RANDOM MARTINGALES AND LOCALIZATION OF MAXIMAL INEQUALITIES

ASSAF NAOR AND TERENCE TAO

ABSTRACT. Let (X, d, μ) be a metric measure space. For $\emptyset \neq R \subseteq (0, \infty)$ consider the Hardy-Littlewood maximal operator

$$M_R f(x) \stackrel{\mathrm{def}}{=} \sup_{r \in R} \frac{1}{\mu(B(x,r))} \int_{B(x,r)} |f| \, d\mu.$$

We show that if there is an $n > 1$ such that one has the "microdoubling condition" $\mu\left(B\left(x, \left(1 + \frac{1}{n}\right) r\right)\right) \lesssim \mu(B(x,r))$ for all $x \in X$ and $r > 0$, then the weak $(1,1)$ norm of M_R has the following localization property:

$$\|M_R\|_{L_1(X) \to L_{1,\infty}(X)} \asymp \sup_{r>0} \left\|M_{R \cap [r,nr]}\right\|_{L_1(X) \to L_{1,\infty}(X)}.$$

An immediate consequence is that if (X, d, μ) is Ahlfors-David n-regular then the weak $(1,1)$ norm of M_R is $\lesssim n \log n$, generalizing a result of Stein and Strömberg [47]. We show that this bound is sharp, by constructing a metric measure space (X, d, μ) that is Ahlfors-David n-regular, for which the weak $(1,1)$ norm of $M_{(0,\infty)}$ is $\gtrsim n \log n$. The localization property of M_R is proved by assigning to each $f \in L_1(X)$ a distribution over *random* martingales for which the associated (random) Doob maximal inequality controls the weak $(1,1)$ inequality for M_R.

CONTENTS

1. Introduction 2
1.1. Microdoubling and the localization theorem 4
1.2. Weak $(1,1)$ norm bounds 5
1.3. Ultrametric approximations: deterministic and random 7
1.4. Lower bounds 8
1.5.1. Adding more hypotheses 9
1.6. The example of the infinite tree 10
2. Doob-type maximal inequalities 11
3. Localization of maximal inequalities 13
4. An argument of E. Lindenstrauss 19
5. The infinite tree 24
6. Sharpness 27
6.1. A preliminary construction 27
6.3. The doubling example 30
6.4. The Ahlfors-David regular example 34
References 39

1. Introduction

A *metric measure space* (X, d, μ) is a separable metric space (X, d), equipped with a Radon measure μ. We assume throughout the non-degeneracy property $0 < \mu(B(x, r)) < \infty$ for all $r > 0$, where $B(x, r) \stackrel{\mathrm{def}}{=} \{y \in X : d(x, y) \leq r\}$. For any locally integrable $f : X \to \mathbb{C}$, we can then define *the Hardy-Littlewood maximal function*

$$Mf(x) \stackrel{\mathrm{def}}{=} \sup_{r>0} \frac{1}{\mu(B(x,r))} \int_{B(x,r)} |f| \, d\mu,$$

which is easily verified to be measurable.

We shall study the *weak* $(1,1)$ *operator norm* of M, defined as usual to be the least quantity $0 \leq \|M\|_{L_1(X)\to L_{1,\infty}(X)} \leq \infty$ for which one has the distributional inequality

$$\|Mf\|_{L_{1,\infty}(X)} \leq \|M\|_{L_1(X)\to L_{1,\infty}(X)} \cdot \|f\|_{L_1(X)} \tag{1}$$

for all $f \in L_1(X)$. Here $L_p(X)$ $(p \geq 1)$ denotes the usual Lebesgue space corresponding to the measure μ, and $L_{p,\infty}(X)$ is the weak L_p norm,

$$\|f\|_{L_{p,\infty}(X)} \stackrel{\mathrm{def}}{=} \sup_{\lambda>0} \lambda \cdot \mu(|f| > \lambda)^{1/p}.$$

Analogously to (1), the *strong* (p,p) operator norm of M is defined as usual to be the least quantity $0 \leq \|M\|_{L_p(X)\to L_p(X)} \leq \infty$ for which

$$\|Mf\|_{L_p(X)} \leq \|M\|_{L_p(X)\to L_p(X)} \cdot \|f\|_{L_p(X)} \tag{2}$$

for all $f \in L_p(X)$.

In most cases of interest it is probably impossible to compute $\|M\|_{L_1(X)\to L_{1,\infty}(X)}$ exactly; notable exceptions to this statement are ultrametric spaces, where the weak $(1,1)$ norm of M equals 1 (we will return to the class of ultrametric spaces presently), and the real line $\mathbb{R}$, equipped with the usual metric and Lebesgue measure, where it was shown by Melas [34] that the weak $(1,1)$ norm of M equals $\frac{11+\sqrt{61}}{12}$ (the case of the strong (p,p) norm of M, $p > 1$, when $X = \mathbb{R}$, remains open, but we refer to [20, 25] for some partial results).

In view of these difficulties, it seems more reasonable to ask for estimates on the asymptotic behavior of the various operator norms of maximal functions. Quite remarkably, despite the wide applicability of maximal inequalities, and significant effort by many researchers, even in the simple case when X is the n-dimensional Hilbert space ℓ_2^n and μ is Lebesgue measure, it is unknown whether or not the weak $(1,1)$ norm of M is bounded independently of the dimension n.

A classical application of the Vitali covering theorem (see for example [17, 46, 21, 27]) shows that for any n-dimensional normed space X, the weak $(1,1)$ and strong (p,p) norms of M grow at most exponentially in n. This was greatly improved by Stein and Strömberg [47] to $\|M\|_{L_1(X)\to L_{1,\infty}(X)} = O(n \log n)$ for a general n-dimensional normed space, and to the slightly better bound $\|M\|_{L_1(\ell_2^n)\to L_{1,\infty}(\ell_2^n)} = O(n)$ for n-dimensional Hilbert space. Until recently, there was no known example of a sequence of n-dimensional normed spaces X_n for which $\|M\|_{L_1(X_n)\to L_{1,\infty}(X_n)}$ tends to ∞ with n. A recent breakthrough of Aldaz [1] showed that when $X_n = \ell_\infty^n$, i.e., $\mathbb{R}^n$ equipped with the ℓ_∞ norm (whose unit ball is an axis parallel cube), $\|M\|_{L_1(X_n)\to L_{1,\infty}(X_n)}$ must tend to ∞ with n; the best known lower bound [3] on $\|M\|_{L_1(\ell_\infty^n)\to L_{1,\infty}(\ell_\infty^n)}$ is $(\log n)^{1-o(1)}$. The best known upper estimate for $\|M\|_{L_1(X)\to L_{1,\infty}(X)}$ when $X = \ell_\infty^n$ remains the Stein-Strömberg $O(n \log n)$ bound.

As partial evidence that when X is the n-dimensional Euclidean space ℓ_2^n, the weak $(1,1)$ norm $\|M\|_{L_1(X)\to L_{1,\infty}(X)}$ might be bounded, we can take Stein's theorem [45] (see also the appendix of [47]) which asserts that in the Euclidean case, for $p > 1$ we have $\|M\|_{L_p(X)\to L_p(X)} \le C(p)$, where $C(p) < \infty$ depends on p but not on n. For general n-dimensional normed spaces, Stein and Strömberg [47] obtained the bound $\|M\|_{L_p(X)\to L_p(X)} \le c(p)n$, while Bourgain [8, 9] and Carbery [13] proved that for any n-dimensional normed space, $\|M\|_{L_p(X)\to L_p(X)} \le C(p) < \infty$ provided $p > \frac{3}{2}$. It is unknown whether or not there is some $1 < p < \frac{3}{2}$ for which there exist n-dimensional normed spaces X_n such that $\|M\|_{L_p(X_n)\to L_p(X_n)}$ is unbounded. This is unknown even for the case of cube averages $X_n = \ell_\infty^n$. It was shown by Bourgain [10] that $\|M\|_{L_p(X)\to L_p(X)} \le C(p,q)$ for all $p > 1$ when $X = \ell_q^n$ and q is an even integer, and this was extended by Müller to $X = \ell_q^n$ for all $1 \le q < \infty$.

A dimension independent bound on $\|M\|_{L_1(\ell_2^n)\to L_{1,\infty}(\ell_2^n)}$ would mean that the classical Euclidean Hardy-Littlewood maximal inequality is in essence an infinite dimensional phenomenon. This statement is not quite true, since there is no "Lebesgue measure" on infinite dimensional Hilbert space, but nevertheless, even Stein's dimension independent bound on $\|M\|_{L_p(\ell_2^n)\to L_p(\ell_2^n)}$, $p > 1$, has interesting infinite dimensional consequences—see for examples Tišer's work [53] on differentiation of integrals with respect to certain Gaussian measures on Hilbert space (provided that the integrand is in L_p for some $p > 1$). Moreover, improved bounds on $\|M\|_{L_1(X)\to L_{1,\infty}(X)}$ are clearly of interest since they would yield improved quantitative estimates in the many known applications of the Hardy-Littlewood maximal inequality. As an example, such bounds are relevant for quantitative variants of Rademacher's differentiation theorem for Lipschitz functions, which are used in results on the bi-Lipschitz distortion of discrete nets (see [11, 15]).

Bounds on $\|M\|_{L_1(X)\to L_{1,\infty}(X)}$ and $\|M\|_{L_p(X)\to L_p(X)}$ have been also intensively investigated for metric measure spaces other than finite dimensional normed spaces. Strong (p,p) bounds for free groups (with counting measure) have been established by Nevo and Stein in [40]. In Section 5 we prove the corresponding weak $(1,1)$ inequality, which is nevertheless not sufficient for the purpose of ergodic theoretical applications as in [40]; see Conjecture 1 below for more information[1]. In the case of the Heisenberg group $\mathbb{H}^{2n+1}$, equipped with either the Carnot-Carathéodory metric or the Koranyi norm (and the underlying measure being the Haar measure), dimension independent strong (p,p) bounds have been obtained by Zienkiewicz [56], and a weak $(1,1)$ bound of $O(n)$ was obtained by Li [30]. It is unclear if these bounds generalize to other nilpotent Lie groups (though perhaps similar methods could apply to certain two step nilpotent Lie groups, by replacing the use of [41] in [56] with the results of [38, 23]).

The main result of the present paper implies a general bound for the weak $(1,1)$ norm of the Hardy-Littlewood maximal function on Ahlfors-David n-regular spaces; a class of metric measure spaces that contains the examples described above as special cases (except for the case of the free group, which is dealt with separately in Section 5). Specifically, assume that

[1] After presenting our work we learned from Michael Cowling that the weak $(1,1)$ inequality for the free group can be also deduced from the work of Rochberg and Taibleson [42]. Our combinatorial proof in Section 5 is different from the proof in [42], though it is similar to the proof in an unpublished manuscript of Cowling, Meda and Setti, which adapts arguments of Strömberg [48] in the case of the hyperbolic space. We thank Michael Cowling and Lewis Bowen for showing us the Cowling-Meda-Setti manuscript.

the metric measure space (X, d, μ) satisfies the growth bounds

$$\forall x \in X \ \forall r > 0, \quad r^n \leq \mu\left(B(x,r)\right) \leq Cr^n, \tag{3}$$

where $n \geq 2$, and C is independent of x, r. Under this assumption, we show that

$$\|M\|_{L_1(X)\to L_{1,\infty}(X)} = O(n \log n), \tag{4}$$

where the implied constant depends only on C. At the same time, we construct for all $n \geq 2$ an Abelian group G_n, equipped with a translation invariant metric d_n and a translation invariant measure μ_n, that satisfies (3) with $C = 81$ [2], yet

$$\|M\|_{L_1(G_n)\to L_{1,\infty}(G_n)} \gtrsim n \log n. \tag{5}$$

We can also ensure that for all $p > 1$ we have

$$\|M\|_{L_p(G_n)\to L_p(G_n)} \lesssim_p 1. \tag{6}$$

Here, and in what follows, we use $X \lesssim Y$, $Y \gtrsim X$ to denote the estimate $X \leq CY$ for some absolute constant C; if we need C to depend on parameters, we indicate this by subscripts, thus $X \lesssim_p Y$ means that $X \leq C_p Y$ for some C_p depending only on p. We shall also use the notation $X \asymp Y$ for $X \lesssim Y \ \wedge \ Y \lesssim X$.

Note that the bound (4) contains the Stein-Strömberg result for n-dimensional normed spaces. It also applies to, say, any translation invariant length metric on nilpotent Lie groups[3]. However, it falls shy (by a logarithmic factor) of the two $O(n)$ results quoted above: for the Euclidean space ℓ_2^n, and the Heisenberg group $\mathbb{H}^{2n+1}$. Our lower bound (5) suggests that in order to improve upon the $O(n \log n)$ bound of Stein and Strömberg, one must genuinely use the underlying geometry of the normed vector space and not just the metric properties, or the L_p theory. For instance, to obtain the bound of $O(n)$ in the case of the Euclidean metric in [47], it was necessary to exploit the relationship between averaging on balls and the Poisson semigroup, in order that the Hopf-Dunford-Schwartz maximal inequality can be used. A similar strategy was used for the Heisenberg group in [30]. This type of relationship does not appear to be available for general norms on $\mathbb{R}^n$.

The results presented above are simple corollaries of a general localization phenomenon for maximal inequalities, which we shall now describe. In fact, for the bound (4) to hold true, we need to assume a condition which is less restrictive than the Ahlfors-David regularity condition (3); in particular it need not hold for all radii r, and it thus also applies to discrete groups of polynomial growth, equipped with the word metric and the counting measure. All of these issues are explained in the following subsection.

1.1. **Microdoubling and the localization theorem.** Let (X, d, μ) be a metric measure space. For $R \subseteq (0, \infty)$ we consider the maximal operator corresponding to radii in R, which is defined by

$$M_R f(x) \stackrel{\mathrm{def}}{=} \sup_{r \in R} \frac{1}{\mu(B(x,r))} \int_{B(x,r)} |f| \, d\mu. \tag{7}$$

Thus, using our previous notation, $M = M_{(0,\infty)}$.

[2] One can modify the argument to make C arbitrarily close to 1, but we will not do so here as it requires more artificial constructions.

[3] It seems likely however that the original Stein-Strömberg argument can be extended to this setting.

We shall say that (X, d, μ) is *n-microdoubling with constant K* if for all $x \in X$ and all $r > 0$ we have

$$\mu\left(B\left(x, \left(1 + \frac{1}{n}\right) r\right)\right) \leq KB(x, r). \tag{8}$$

The case $n = 1$ in (8) is the classical *K-doubling* condition

$$\forall x \in X\ \forall r > 0, \quad \mu\left(B\left(x, 2r\right)\right) \leq KB(x, r). \tag{9}$$

Note that (8) follows from the Ahlfors-David n-regularity condition (3), with $K = eC$. The microdoubling property appeared in various guises in the literature; for example, it follows from a lemma of Colding and Minicozzi [18] (see also Proposition 6.12 in [14]) that if (X, d, μ) is a K-doubling *length* space, then it is also n-microdoubling with constant $O(1)$, where $n = e^{K^{O(1)}}$. We note in passing that this exponential dependence on K is necessary, as exhibited by the interval $X = [1, N]$, with the metric inherited from $\mathbb{R}$, and the measure whose density is $\varphi(x) = \frac{1}{x}$; the doubling constant for this length space is of order $\log N$, but it can only be n-microdoubling with n a power of N.

Our main result is the following localization theorem for maximal inequalities on microdoubling spaces. It deals, for any $1 \leq p < \infty$, with the weak (p, p) norm of M_R, defined as the optimal number $\|M_R\|_{L_p(X) \to L_{p,\infty}(X)}$ for which the distributional inequality

$$\mu\left(M_R f > \lambda\right) \leq \frac{\|M_R\|^p_{L_p(X) \to L_{p,\infty}(X)}}{\lambda^p} \|f\|^p_{L_p(X)}$$

holds for all $f \in L_p(X)$ and $\lambda > 0$.

Theorem 1.1 (Localisation)**.** *Fix $n \geq 1$ and $K \geq 5$. Let (X, d, μ) be a metric measure space satisfying the microdoubling condition (8). Fix $\emptyset \neq R \subseteq (0, \infty)$ and $p \geq 1$. Then we have*

$$\|M_R\|_{L_p(X) \to L_{p,\infty}(X)} \lesssim K + \left(1 + \frac{\log \log K}{1 + \log n}\right)^{1/p} \sup_{r > 0} \left\|M_{R \cap [r, nr]}\right\|_{L_p(X) \to L_{p,\infty}(X)}. \tag{10}$$

Remark 1.1. In the converse direction, one trivially has

$$\|M_R\|_{L_p(X) \to L_{p,\infty}(X)} \geq \sup_{r > 0} \left\|M_{R \cap [r, nr]}\right\|_{L_p(X) \to L_{p,\infty}(X)}.$$

Note that the term $\frac{\log \log K}{1 + \log n}$ in (10) is always at most $\log \log K$. Thus when K is independent of n, up to constants, in order to establish a weak (p, p) maximal inequality for spaces obeying (8), it suffices to do so for scales localized to an interval $[r, nr]$. In many cases (e.g. finite-dimensional normed vector spaces) we can also rescale to $r = 1$.

1.2. **Weak $(1, 1)$ norm bounds.** To deduce some corollaries of Theorem 1.1, fix an integer $m \in \mathbb{N}$, and note that for all $f \in L_p(X)$ and $r, \lambda > 0$ we have,

$$\begin{aligned}\mu\left(M_{R \cap [r, nr]} f > \lambda\right) &= \mu\left(\max_{0 \leq j \leq m-1} M_{R \cap \left[rn^{j/m}, rn^{(j+1)/m}\right]} f > \lambda\right) \\ &\leq \sum_{j=0}^{m-1} \mu\left(M_{R \cap \left[rn^{j/m}, rn^{(j+1)/m}\right]} f > \lambda\right) \leq m \max_{0 \leq j \leq m-1} \mu\left(M_{R \cap \left[rn^{j/m}, rn^{(j+1)/m}\right]} f > \lambda\right).\end{aligned}$$

Thus, under the assumptions of Theorem 1.1 (and specializing to $p = 1$), we have for every $m \in \mathbb{N}$,

$$\|M_R\|_{L_1(X)\to L_{1,\infty}(X)} \lesssim K + m\left(1 + \frac{\log\log K}{1+\log n}\right)\sup_{r>0}\left\|M_{R\cap\left[r,n^{1/m}r\right]}\right\|_{L_1(X)\to L_{1,\infty}(X)}. \tag{11}$$

Note that for $m \geq 2n\log n$ we have $n^{1/m} \leq 1 + \frac{1}{n}$, and hence for all $r > 0$,

$$M_{R\cap\left[r,n^{1/m}r\right]}f \leq \frac{1}{\mu(B\left(x,r\right))}\int_{B\left(x,\left(1+\frac{1}{n}\right)r\right)}|f|d\mu \overset{(8)}{\leq} KA_{1+\frac{1}{n}}f, \tag{12}$$

where A_r is the averaging operator:

$$A_r f(x) \overset{\mathrm{def}}{=} \frac{1}{\mu(B(x,r))}\int_{B(x,r)}|f|\,d\mu. \tag{13}$$

Under some mild uniformity assumption on μ, the strong $(1,1)$ norm of A_r is bounded for all $r > 0$. For example, if $\mu(B(x,r))$ does not depend on x (as is the case for invariant metrics and measures on groups), then a simple application of Fubini's theorem shows that $\|A_r\|_{L_1(X)\to L_1(X)} \leq 1$. In fact, if we knew that $\mu(B(x,r)) \leq K\mu(B(y,r))$ for all $x \in X$ and $y \in B(x,r)$ (which is a trivial consequence of the Ahlfors-David regularity condition (3)), then we would have by the same reasoning $\|A_r\|_{L_1(X)\to L_1(X)} \leq K$. An elegant way to combine this uniformity condition with the microdoubling condition (8), is to impose the following condition, which we call *strong n-microdoubling with constant K*:

$$\forall x \in X\ \forall r > 0\ \forall y \in B(x,r), \quad \mu\left(B\left(y,\left(1+\frac{1}{n}\right)r\right)\right) \leq KB(x,r). \tag{14}$$

Thus, by a combination of (11) and (12), we see that if (X,d,μ) satisfies (14), then $\|M\|_{L_1(X)\to L_{1,\infty}(X)} \lesssim_K n\log n$. Similarly, if $R\cap\left[r,n^{1/m}r\right]$ contains at most one point for all $r > 0$, then $\|M_R\|_{L_1(X)\to L_{1,\infty}(X)} \lesssim_K m$. This happens in particular if

$$R = 2^{\mathbb{Z}} = \left\{2^k :\ k \in \mathbb{Z}\right\},$$

and $m \asymp \log n$, proving the following corollary:

Corollary 1.2. *Fix $n \geq 1$ and $K \geq 5$. Let (X,d,μ) be a metric measure space satisfying the strong n-microdoubling condition (14). Then*

$$\|M\|_{L_1(X)\to L_{1,\infty}(X)} \lesssim_K n\log n, \tag{15}$$

$$\|M_{2^{\mathbb{Z}}}\|_{L_1(X)\to L_{1,\infty}(X)} \lesssim_K \log n. \tag{16}$$

The lacunary maximal function $M_{2^{\mathbb{Z}}}$ was previously studied for n-dimensional normed spaces by Bourgain in [9], where he proved that its strong (p,p) norm is bounded by a dimension independent constant $C_p < \infty$ (recall that for the non-lacunary maximal function this is only known for $p > \frac{3}{2}$). The logarithmic upper bound (16) on the weak $(1,1)$ norm of the lacunary maximal function when X is an n-dimensional normed space was proved by Menárguez and Soria in [35].

In section 4 we present a different approach to the proof of Corollary 1.2, following an argument of E. Lindenstrauss [31]. While it gives slightly weaker results, and does not yield the localization theorem, this approach is of independent interest. Moreover, Lindenstrauss' approach is based on a beautiful randomization of the Vitali covering argument, and as

such complements our approach to Theorem 1.1, which is based on a random partitioning method that originated in theoretical computer science and combinatorics (an overview of our technique is contained in Section 1.3). The maximal functions considered in [31] arose when taking averages over Følner sequences of an amenable group action on a measure space, and were thus not directly connected to the metric questions that are studied in the present paper. Nevertheless we consider the arguments in Section 4 to be essentially the same as those in [31]. We thank Raanan Schul for pointing out how the maximal inequality of E. Lindenstrauss implies the Hardy-Littlewood maximal inequality under strong microdoubling.

1.3. **Ultrametric approximations: deterministic and random.** Doob's classical maximal inequality for martingales (see Section 2) is perhaps the simplest and most versatile maximal inequality for which the weak $(1,1)$ norm is known exactly (and is equal to 1). Our proof of Theorem 1.1 relates the weak $(1,1)$ inequality for M to the maximal inequality for martingales, by allowing the martingale itself to be a random object. We show that while the weak $(1,1)$ inequality is not itself a martingale inequality, it is possible to associate to each $f \in L_1(X)$ a distribution over *random martingales*. These random martingales stochastically approximate Mf, in the sense that we can write down a variant of Doob's inequality for each of them, which, under the microdoubling assumption, in expectation yields theorem 1.1. The details are presented in Section 3.

An alternative interpretation of Doob's maximal inequality is that if (X,d,μ) is a metric measure space, and if in addition d is an ultrametric, i.e., $d(x,y) \le \max\{d(x,z), d(z,y)\}$ for all $x,y,z \in X$, then $\|M\|_{L_1(X)\to L_{1,\infty}(X)} \le 1$. Indeed, restrict for simplicity to the case of a finite ultrametric, in which case we obtain an induced hierarchical family of partitions of X into balls, where each ball at a given "level" is the union of balls of smaller radii at the next "level". This picture immediately shows that by considering the averages of f on smaller and smaller balls, in the ultrametric case we can reduce the weak $(1,1)$ inequality for Mf to Doob's maximal inequality.

Of course, not every metric is an ultrametric, or even close to an ultrametric. Nevertheless, over the previous two decades, researchers in combinatorics and computer science developed methods to associate to a general metric space (X,d) a distribution over *random ultrametrics* ρ on X, which dominate d and sufficiently approximate it in various senses (depending on the application at hand). Such methods are often also called "random partitioning methods", in reference to the hierarchical (tree) structure of ultrametrics. This approach originated in the pioneering works of Linial and Saks [32] and Alon, Karp, Peleg and West [2], and has been substantially developed and refined by Bartal [4, 5]. Important contributions of Calinescu, Karloff and Rabani [12] and Fakcharoenphol, Rao and Talwar [22] resulted in a sharp form of "Bartal's random tree method", and our work builds on these ideas. In [36, 37] such random ultrametrics were used in order to prove maximal-type inequalities of a very different nature (motivated by embedding problems, as ultrametrics are isometric to subsets of Hilbert space [29]); these results also served as some inspiration for our work.

One should mention here that the idea of relating metrics to ultrametric models is, of course, standard. Hierarchical partitioning schemes are ubiquitous in analysis and geometry (see the discussion of Calderón-Zygmund decompositions in [45], or, say, Christ's cube construction in [16]). Proving maximal inequalities by considering certain Hierarchical partitions is extremely natural; a striking example of this type is Talagrand's majorizing measure

theorem [49], which deals with sharp maximal inequalities for Gaussian processes via a construction of special ultrametrics (the ultrametric approach is explicit in [49], and has an alternative later description [50] via the so called "generic chaining"; see also [26]). Explicit uses of random coverings and partitions in the context of purely analytic problems occurred in E. Lindenstrauss' aforementioned randomization of the Vitali covering argument for the purpose of pointwise theorems for amenable groups [31], and in the work of Nazarov, Treil and Volberg [39] on $T(b)$ theorems on non-homogeneous spaces. See also [28] for applications to extensions of Lipschitz functions.

1.4. **Lower bounds.** A standard application of the Vitali covering argument (see e.g. [46] or [52]) yields the inequality

$$\left\|\tilde{M}f\right\|_{L_{1,\infty}(X)} \le \|f\|_{L_1(X)}, \tag{17}$$

where $\tilde{M}f$ is the modified Hardy-Littlewood maximal operator

$$Mf(x) \stackrel{\text{def}}{=} \sup_{r>0} \frac{1}{\mu(B(x,r,r))} \int_{B(x,r)} |f|\, d\mu,$$

and $B(x,r) \subseteq B(x,r,r) \subseteq B(x,2r)$ is the enlarged ball

$$B(x,r,r) \stackrel{\text{def}}{=} \bigcup_{y\in B(x,r)} B(y,r) = \{z \in X : d(x,y), d(y,z) \le r \text{ for some } y \in X\}.$$

In particular, if we have the doubling condition (9), then

$$\|M\|_{L_1(X)\to L_{1,\infty}(X)} \le K. \tag{18}$$

The factor 2 in (9) cannot be replaced by any smaller number while still retaining linear behavior in terms of K of the weak $(1,1)$ operator norm; see [43].

In the absence of any further assumptions on the metric measure space, the bound (18) is close to sharp:

Proposition 1.5 (The star counterexample)**.** *Fix $K \ge 1$. Then there exists a metric measure space obeying (9) with*

$$\|M\|_{L_1(X)\to L_{1,\infty}(X)} \ge \lfloor K \rfloor - 1.$$

Proof. Without loss of generality we may take K to be an integer. Let X be the "star" graph formed by connecting one "hub" vertex v_0 to $(K-1)^2$ other "spoke" vertices $v_1, \ldots, v_{(K-1)^2}$, with the usual graph metric (thus $d(v_0, v_i) = 1$ and $d(v_i, v_j) = 2$ for all distinct $i, j \in \{1, \ldots, (K-1)^2\}$). Let μ be the measure which assigns the mass $K-1$ to v_0 and mass 1 to all other vertices; one easily verifies that (9) holds. Let $f \in L_1(X)$ be the function which equals 1 on v_0 and vanishes elsewhere. Then one easily verifies that $\|f\|_{L_1(X)} = K-1$, that $\mu(X) = K(K-1)$, and that $Mf(x) \ge \frac{K-1}{K}$ for all $x \in X$, and the claim follows. $\square$

Remark 1.2. One can achieve a similar effect in a high-dimensional Euclidean space $\mathbb{R}^n$. If we let $X = \{0, e_1, \ldots, e_n\}$ be the origin and standard basis with the usual Euclidean metric and counting measure, then (9) holds with $K \stackrel{\text{def}}{=} n+1$, while if we let f be the indicator function of 0, then $Mf(x) \ge \frac{1}{2}$ for all $x \in X$, and so $\|M\|_{L_1(X)\to L_{1,\infty}(X)} \ge \frac{n+1}{2} = \frac{K}{2}$. A more sophisticated version of this example was observed in [44]: if we take X to be the origin 0, together with a maximal 1.01-separated (say) subset of the sphere S^{d-1}, then (9) holds for $K = |X| \ge C^n$ for some absolute constant $C > 1$, but $\|M\|_{L_1(X)\to L_{1,\infty}(X)} \ge \frac{K}{2}$ by the

same argument as before. In particular this shows that the Hardy-Littlewood weak $(1,1)$ operator norm (as well as the L_p operator norm for any fixed $1 < p < \infty$) for measures in $\mathbb{R}^n$ can grow exponentially in the dimension n. In the converse direction, a well-known application of the Besicovitch covering lemma [6, 7] shows that $\|M\|_{L_1(X)\to L_{1,\infty}(X)} \le C^n$ for some absolute constant C whenever X is a subset of $\mathbb{R}^n$ with the Euclidean metric, and μ is an arbitrary Radon measure. In particular, as observed in [44], this shows that the constants in the Besicovitch covering lemma must grow exponentially in the dimension (see also [24]).

1.5.1. *Adding more hypotheses.* Despite the example in Proposition 1.5, we know due to Corollary 1.2 that in many cases the bound (18) can be significantly improved. In particular, a more meaningful variant of Proposition 1.5 would be if we also impose the natural uniformity condition that $\mu(B(x,r))$ is independent of $x \in X$. As discussed in Section 1.2, this immediately implies that the averaging operators A_r given in (13) are now contractions on $L_1(X)$. Thus in order for the weak $(1,1)$ operator norm to be large, one needs to have contributions to the set $\{Mf > \lambda\}$ from several scales r, rather than just a single scale as in Proposition 1.5.

Another hypothesis that one can add, in order to make a potential counter-example more meaningful, is that the maximal operator M is already of strong-type (p,p) for all $1 < p \le \infty$, as we know to be the case for $X = \ell_2^n$, due to Stein's theorem [45]. Finally, we can make the task of bounding the maximal operator easier by replacing M with the lacunary maximal operator $M_{2^{\mathbb{Z}}}$.

Our first main construction shows that even with all of these additional hypotheses and simplifications, we still cannot improve significantly upon (18).

Theorem 1.3 (Doubling example)**.** *Let $K \ge 1$. Then there exists a metric measure space (X,d,μ) with X an Abelian group and d,μ translation-invariant, such that the doubling condition (9) holds, and $\|M\|_{L_p(X)\to L_p(X)} \lesssim_p 1$ holds for all $1 < p \le \infty$ (with the implied constant independent of K), but such that*

$$\|M_{2^{\mathbb{Z}}}\|_{L_1(X)\to L_{1,\infty}(X)} \ge \frac{K}{48}. \tag{19}$$

We prove this theorem in Section 6.3. The basic idea is to first build a maximal operator not arising from a metric measure space which is of strong type (p,p) but not of weak type $(1,1)$, and then take an appropriate "tensor product" of this operator with a martingale type operator to obtain a new operator which is essentially a lacunary maximal operator associated to a metric measure space. The constant 48 in (19) can of course be improved, but we will not seek to optimize it here.

As stated earlier, we also construct an example of a metric measure space that shows that Corollary 1.2 is sharp even under the stronger Ahlfors-David regularity condition (3).

Theorem 1.4 (Ahlfors-David regular example)**.** *Assume that $n \ge 2$. Then there exists an Abelian group G, with invariant measure μ and an invariant metric d, obeying the Ahlfors-David n-regularity condition (3) with $K = 81$, such that*

$$\|M\|_{L_1(G)\to L_{1,\infty}(G)} \gtrsim n \log n, \tag{20}$$

and

$$\|M_{2^{\mathbb{Z}}}\|_{L_1(G)\to L_{1,\infty}(G)} \gtrsim \log n. \tag{21}$$

Furthermore we have

$$\|M\|_{L_p(G)\to L_p(G)} \lesssim_p 1 \tag{22}$$

for all $1 < p \leq \infty$.

1.6. **The example of the infinite tree.** The above examples seem to indicate that the weak $(1,1)$ behavior of the Hardy-Littlewood maximal function can deteriorate substantially when the doubling constant is large, even when assuming good L_p bounds, as well as uniformity assumptions on the measure of balls. Nevertheless, there are some interesting examples of metric measure spaces with very poor (or non-existent) doubling properties, for which one still has a weak $(1,1)$ bound. We give just one example of this phenomenon, namely the infinite regular tree.

Theorem 1.5 (Hardy-Littlewood inequality for the infinite tree)**.** *Fix an integer* $k \geq 2$*, and let* T *be the infinite rooted* k*-ary tree, with the usual graph metric* d *and counting measure* μ*. Then we have*

$$\|M\|_{L_1(T)\to L_{1,\infty}(T)} \lesssim 1$$

(Thus the implied constant is independent of the degree k*.)*

We prove this theorem in Section 5. We remark that the L_p boundedness of this maximal function for $p > 1$ was essentially established by Nevo and Stein in [40]. The argument here proceeds very differently from the usual covering type arguments, which are totally unavailable here due to the utter lack of doubling for this tree. Instead, we use a more combinatorial argument taking advantage of the "expander" or "non-amenability" properties of this tree, which roughly asserts that any given finite subset of the tree must have large boundaries at every distance scale.

When k is odd, T is almost[4] identifiable with the free group on $\frac{k+1}{2}$ generators. The above theorem then suggests that a maximal ergodic theorem in L_1 should be available for ergodic actions of free groups on measure-preserving systems (the analogous L_p maximal theorems for $p > 1$ being established in [40]). However, the non-amenability of the free group prevents one from applying standard arguments to transfer Theorem 1.5 to this setting (indeed, our proof of Theorem 1.5 will rely heavily on this non-amenability). Thus the following conjecture remains open:

Conjecture 1. *Let* F *be a finitely generated free group, and let* $w \mapsto T_w$ *be an ergodic action of* F *on a probability space* $(X, \mathscr{B}, \mu)$*. Then*

$$\left\| \sup_{n\geq 1} \frac{1}{|B(\mathrm{id}, n)|} \sum_{w\in B(\mathrm{id},n)} |T_w f| \right\|_{L_{1,\infty}(X)} \lesssim \|f\|_{L_1(X)}$$

for all $f \in L_1(X)$*, where* $B(\mathrm{id}, n)$ *is the collection of words in* F *of length less than* n*.*

We remark that by applying the pointwise convergence theorems in [40] and a standard density argument, Conjecture 1 would imply the pointwise convergence result

$$\lim_{n\to\infty} \frac{1}{|B(\mathrm{id}, n)|} \sum_{w\in B(\mathrm{id},n)} T_w f(x) = \int_X f \, d\mu$$

[4]More precisely, one needs to enlarge the tree at the root to have $k+1$ descendants instead of k. But one can easily check that this change only affects the weak $(1,1)$ norm of the maximal function by a constant at worst.

for all $f \in L_1(X)$ and almost every $x \in X$. This result is currently known for $f \in L_p(X)$ for $p > 1$, due to [40].

Acknowledgements. We thank Raanan Schul for pointing out that the Lindenstrauss maximal inequality implies the Hardy-Littlewood maximal inequality under strong microdoubling, and Zubin Guatam for explaining the proof of the Lindenstrauss maximal inequality. A. N. was supported in part by NSF grants CCF-0635078 and CCF-0832795, BSF grant 2006009, and the Packard Foundation. T. T. was supported by a grant from the MacArthur foundation, by NSF grant DMS-0649473, and by the NSF Waterman award.

2. Doob-type maximal inequalities

Let (X, d, μ) be a metric measure space with $\mu(X) < \infty$ (more generally, the arguments below extend to the σ-finite case). If $\mathscr{F}$ is a σ-algebra of measurable sets in X, we let $L_p(\mathscr{F})$ denote the space of $L_p(X)$ functions which are $\mathscr{F}$-measurable. The orthogonal projection from $L_2(X)$ to the closed subspace $L_2(\mathscr{F})$ will be denoted $f \mapsto \mathbb{E}(f|\mathscr{F})$, and as is well known it extends to a contraction on $L_p(X)$ for all $1 \le p \le \infty$. The following important inequality of Doob is classical (see [19, 21]).

Proposition 2.1 (Doob's maximal inequality). *Let $\mathscr{F}_0 \subseteq \mathscr{F}_1 \subseteq \mathscr{F}_2 \subseteq \cdots$ be an increasing sequence of σ-algebras. Then we have*

$$f \in L_1(X) \implies \left\| \sup_{k \ge 0} \left|\mathbb{E}(f|\mathscr{F}_k)\right| \right\|_{L_{1,\infty}(X)} \le \|f\|_{L_1(X)},$$

and for $1 < p \le \infty$,

$$f \in L_p(X) \implies \left\| \sup_{k \ge 0} \left|\mathbb{E}(f|\mathscr{F}_k)\right| \right\|_{L_p(X)} \le \frac{p}{p-1} \|f\|_{L_p(X)}.$$

We now establish a variant of this inequality, in which the expectations $\mathbb{E}(f|\mathscr{F}_k)$ are replaced by more general sublinear operators.

Theorem 2.1 (Modified Doob's inequality). *Let $\mathscr{F}_0 \subseteq \mathscr{F}_1 \subseteq \mathscr{F}_2 \subseteq \cdots$ be an increasing sequence of σ-algebras and fix $1 \le p < \infty$. For each $k \in \mathbb{N}$ let M_k be a sublinear operator[5] defined on $L_p(X) + L_\infty(X)$ such that we have the bounds*

$$f \in L_p(X) \implies \|M_k f\|_{L_{p,\infty}(X)} \le A \|f\|_{L_p(X)}, \tag{23}$$

and

$$f \in L_\infty(X) \implies \|M_k f\|_{L_\infty(X)} \le B \left\|\mathbb{E}\left(|f|\middle|\mathscr{F}_k\right)\right\|_{L_\infty(X)}. \tag{24}$$

Suppose also that we have the localization property

$$f \in L_p(X) + L_\infty(X) \ \wedge \ E_k \in \mathscr{F}_k \implies \mathbf{1}_{E_k} M_{k+1} f = M_{k+1}\left(\mathbf{1}_{E_k} f\right). \tag{25}$$

Then we have

$$\left\| \sup_{k \ge 0} |M_k f| \right\|_{L_{p,\infty}(X)} \le \left((2A)^p + (2B)^p\right)^{1/p} \|f\|_{L_p(X)}$$

for all $f \in L_p(X)$.

---

[5] By this we mean that $|M_k(f+g)| \le |M_k(f)| + |M_k(g)|$ and $|M_k(cf)| = |c| \cdot |M_k f|$ for all functions f, g in the domain of M_k and all constants $c \in \mathbb{R}$.

Remark 2.1. Observe that the properties (24), (25) (with $B = 1$) are satisfied by the projection operator $M_{k+1}f \stackrel{\text{def}}{=} \mathbb{E}(f|\mathscr{F})$ whenever $\mathscr{F}_k \subseteq \mathscr{F} \subseteq \mathscr{F}_{k+1}$. Thus (24), (25) can be viewed together as a kind of assertion that M_{k+1} lies "between" $\mathscr{F}_k$ and $\mathscr{F}_{k+1}$ in some sense.

Proof. By monotone convergence we may restrict the supremum over $k \geq 0$ to a finite range, say $0 \leq k \leq K$ for some finite $K \in \mathbb{N}$. We can then assume without loss of generality that $\mathscr{F}_k$ is the trivial algebra $\{\emptyset, X\}$ for all $k < 0$. By homogeneity it suffices to show that

$$f \in L_p(X) \implies \mu\left(\sup_{0\leq k\leq K} |M_k f| > 1\right) \leq ((2A)^p + (2B)^p)\int_X |f|^p \, d\mu. \tag{26}$$

Fix $f \in L_p(X)$ and note that Doob's maximal inequality implies that

$$\mu\left(\sup_{0\leq k\leq K} \mathbb{E}\left(|f|\big|\mathscr{F}_k\right) \geq \frac{1}{2B}\right) \leq \mu\left(\sup_{0\leq k\leq K} \mathbb{E}\left(|f|^p\big|\mathscr{F}_k\right) \geq \frac{1}{(2B)^p}\right) \leq (2B)^p \int_X |f|^p \, d\mu.$$

Thus in order to prove (26) it will suffice to show that

$$\mu\left(\left\{\sup_{0\leq k\leq K} |M_k f| > 1\right\} \setminus \left\{\sup_{0\leq k\leq K} \mathbb{E}\left(|f|\big|\mathscr{F}_k\right) \geq \frac{1}{2B}\right\}\right) \leq (2A)^p \int_X |f|^p \, d\mu. \tag{27}$$

Consider the inclusion

$$\begin{aligned}\left\{\sup_{0\leq k\leq K} |M_k f| > 1\right\} &\setminus \left\{\sup_{0\leq k\leq K} \mathbb{E}\left(|f|\big|\mathscr{F}_k\right) \geq \frac{1}{2B}\right\} \\ &\subseteq \bigcup_{k=0}^{K} \left\{|M_k f| > 1 \ \wedge \ \sup_{0\leq j<k} \mathbb{E}\left(|f|\big|\mathscr{F}_j\right) < \frac{1}{2B}\right\}. \end{aligned}\tag{28}$$

Therefore, if we introduce the sets

$$A_k \stackrel{\text{def}}{=} X \setminus \bigcup_{0\leq j<k} \left\{\mathbb{E}\left(|f|\big|\mathscr{F}_j\right) \geq \frac{1}{2B}\right\},$$

and

$$\Omega_k \stackrel{\text{def}}{=} \left\{\mathbb{E}\left(|f|\big|\mathscr{F}_k\right) \geq \frac{1}{2B}\right\} \cap A_k.$$

Then $A_k \in \mathscr{F}_{k-1}$, the sets Ω_k are disjoint, and using (25) we see that (28) implies the inclusion

$$\begin{aligned}\left\{\sup_{0\leq k\leq K} |M_k f| > 1\right\} \setminus \left\{\sup_{0\leq k\leq K} \mathbb{E}\left(|f|\big|\mathscr{F}_k\right) \geq \frac{1}{2B}\right\} &\subseteq \bigcup_{k=0}^{K} \{|\mathbf{1}_{A_k} M_k f| > 1\} \\ &= \bigcup_{k=0}^{K} \{|M_k(\mathbf{1}_{A_k} f)| > 1\}. \end{aligned}\tag{29}$$

On the other hand, from (24) we have

$$\begin{aligned}\left\|M_k(f\mathbf{1}_{A_k\setminus\Omega_k})\right\|_{L_\infty(X)} \leq B\left\|\mathbb{E}\left(|f|\mathbf{1}_{A_k\setminus\Omega_k}|\mathscr{F}_k\right)\right\|_{L_\infty(X)} &= B\left\|\mathbb{E}\left(|f|\big|\mathscr{F}_k\right)\mathbf{1}_{A_k\setminus\Omega_k}\right\|_{L_\infty(X)} \\ &\leq B\cdot\frac{1}{2B} = \frac{1}{2}. \end{aligned}$$

Hence by the sublinearity of M_k we have the following inclusion (up to sets of measure zero):

$$\{|M_k(f\mathbf{1}_{A_k})| > 1\} \subseteq \left\{|M_k(f\mathbf{1}_{\Omega_k})| > \frac{1}{2}\right\}. \tag{30}$$

Combining (29) with (30) and the assumption (23), we obtain

$$\mu\left(\left\{\sup_{0\le k\le K} |M_k f| > 1\right\} \setminus \left\{\sup_{0\le k\le K} \mathbb{E}\left(|f|\middle|\mathscr{F}_k\right) \ge \frac{1}{2B}\right\}\right) \le \sum_{k=0}^{K} \mu\left(|M_k(f\mathbf{1}_{\Omega_k})| > \frac{1}{2}\right)$$

$$\le \sum_{k=0}^{K} (2A)^p \int_{\Omega_k} |f|^p \, d\mu = (2A)^p \int_{\bigcup_{k=0}^K \Omega_k} |f|^p d\mu \le (2A)^p \int_X |f|^p \, d\mu.$$

This is precisely the estimate (27), as desired. □

3. Localization of maximal inequalities

Let (X, d, μ) be a bounded metric measure space. Given a partition $\mathscr{P}$ of X and $x \in X$, we denote by $\mathscr{P}(x)$ the unique element of $\mathscr{P}$ containing X. We shall say that a sequence $\{\mathscr{P}_k\}_{k=0}^{\infty}$ of partitions of X is a *partition tree* if the following conditions hold true:

- $\mathscr{P}_0$ is the trivial partition $\{X\}$.
- For every $x \in X$ and $k \in \{0\} \cup \mathbb{N}$ we have
$$\operatorname{diam}(\mathscr{P}_k(x)) \le \frac{\operatorname{diam}(X)}{2^k}. \tag{31}$$
- For every $k \in \{0\} \cup \mathbb{N}$ the partition $\mathscr{P}_{k+1}$ is a refinement of the partition $\mathscr{P}_k$, i.e., for every $x \in X$ we have $\mathscr{P}_{k+1}(x) \subseteq \mathscr{P}_k(x)$.

For $\beta > 0$, a probability distribution Pr over partition trees $\{\mathscr{P}_k\}_{k=0}^{\infty}$ is said to be *β-padded* if for every $x \in X$ and every $k \in \mathbb{N}$,

$$\Pr\left[B\left(x, \frac{\beta \operatorname{diam}(X)}{2^k}\right) \subseteq \mathscr{P}_k(x)\right] \ge \frac{1}{2}. \tag{32}$$

Note that (32) has the following simple consequence, which we will use later: for every measurable set $\Omega \subseteq X$ denote

$$\Omega_\beta^{\mathrm{pad}(k)} \stackrel{\mathrm{def}}{=} \left\{x \in \Omega : \; B\left(x, \frac{\beta \operatorname{diam}(X)}{2^k}\right) \subseteq \mathscr{P}_k(x)\right\}. \tag{33}$$

Thus $\Omega_\beta^{\mathrm{pad}(k)}$ is a random subset of Ω. By Fubini's theorem we have:

$$\mathbb{E}\left[\mu\left(\Omega_\beta^{\mathrm{pad}(k)}\right)\right] = \int_\Omega \Pr\left[B\left(x, \frac{\beta \operatorname{diam}(X)}{2^k}\right) \subseteq \mathscr{P}_k(x)\right] d\mu(x) \stackrel{(32)}{\ge} \frac{\mu(\Omega)}{2}. \tag{34}$$

Remark 3.1. In the definitions above we implicitly made the assumptions that certain events are measurable in the appropriate measure spaces. Namely, for (32) we need the event $\left\{B\left(x, \frac{\beta \operatorname{diam}(X)}{2^k}\right) \subseteq \mathscr{P}_k(x)\right\}$ to be Pr-measurable for every $x \in X$ and $k \in \{0\} \cup \mathbb{N}$, and for (34) we need the event $\left\{(x, \{\mathscr{P}_k\}_{k=0}^{\infty}) : \; x \in \Omega \;\wedge\; B\left(x, \frac{\beta \operatorname{diam}(X)}{2^k}\right) \subseteq \mathscr{P}_k(x)\right\}$ to be measurable with respect to $\mu \times \Pr$ for all $k \in \{0\} \cup \mathbb{N}$. These assumptions will be trivially satisfied in the concrete constructions below.

Remark 3.2. In the above definitions we made some arbitrary choices: the factor $\frac{1}{2^k}$ in (31) can be taken to be some other factor $r_k > 0$, and the $\frac{1}{2}$ lower bound on the probability in (32) can be taken to be some other probability p_k. Since we will not use these additional degrees of freedom here, we chose not to mention them for the sake of simplifying notation. But, the arguments below can be easily carried out in greater generality, which might be useful for future applications of these notions.

The following lemma deals with the existence of padded random partition trees on microdoubling metric measure spaces. The argument is similar to the proof of Theorem 3.17 in [28], which is based on ideas from the theoretical computer science literature [12, 22]. The last part of the argument is in the spirit of the proof of the main padding inequality in [37].

Lemma 3.1. *Fix $n \geq 1$ and $K \geq 5$. Let (X, d, μ) be a separable bounded metric measure space which satisfies (8). Then X admits a $\frac{1}{16n \log K}$-padded probability distribution over partition trees.*

Remark 3.3. Let (X, d) is a separable complete and bounded metric space which is doubling with constant λ, i.e., every ball in X can be covered by at most λ balls of half the radius. It is a classical fact, due to Vol′berg and Konyagin [54] in the case of compact spaces, and Luukkainen and Saksman [33] in the case of general complete spaces (see also [55] and chapter 13 in [27]), that X admits a non-degenerate measure μ which is doubling with constant λ^2 (the power 2 can be replaced here by any power bigger than 1). Thus the conclusion of Lemma 3.1 holds in this case with $n = 1$ and $K = \lambda^2$.

Proof of Lemma 3.1. By rescaling the metric we may assume without loss of generality that $\operatorname{diam}(X) = 1$. Since X is bounded, $\mu(X) < \infty$, and we may therefore normalize μ to be a probability measure. Let $x_1, x_2, x_3, \ldots$ be points chosen uniformly and independently at random from X according to the measure μ, i.e., $(x_1, x_2, \ldots)$ is distributed according to the probability measure $\mu^{\otimes \aleph_0}$. For each k let r_k be a random variable that is distributed uniformly on the interval $\left[2^{-k-2}, 2^{-k-1}\right]$. We assume that $r_1, r_2, \ldots$ are independent. Let Pr denote the joint distribution of $(x_1, x_2, \ldots), (r_1, r_2, \ldots)$.

For every $k \in \mathbb{N}$ define a random variable $j_k : X \to \mathbb{N} \cup \{\infty\}$ by

$$j_k(x) \stackrel{\mathrm{def}}{=} \inf \left\{ j \in \mathbb{N} \cup \{\infty\} : \; d(x, x_j) \leq r_k \right\}.$$

Note that $j_k(x)$ is almost surely finite for every $x \in X$, since each x_j has positive probability of falling into $B(x, r_k) \supseteq B\left(x, 2^{-k-2}\right)$ (see the argument in [28] for more details). Since X is separable, it follows that the event $\bigcup_{x \in X} \bigcup_{k=1}^{\infty} \{j_k(x) < \infty\}$ has probability 1. From now on we will condition on this event.

For every $k \in \mathbb{N}$ and $\ell_1, \ldots, \ell_k \in \mathbb{N}$ define

$$P(\ell_1, \ldots, \ell_k) \stackrel{\mathrm{def}}{=} \left\{ x \in X : \; j_1(x) = \ell_1, \ldots, j_k(x) = \ell_k \right\}.$$

Then $\mathscr{P}_k \stackrel{\mathrm{def}}{=} \left\{ P(\ell_1, \ldots, \ell_k) : \; \ell_1, \ldots, \ell_k \in \mathbb{N} \right\}$ is a partition of X. By definition

$$P(\ell_1, \ldots, \ell_k) \subseteq B(x_{\ell_k}, r_k) \subseteq B\left(x_{\ell_k}, 2^{-k-1}\right),$$

and for all $k \in \mathbb{N}$,

$$P(\ell_1, \ldots, \ell_k, \ell_{k+1}) \subseteq P(\ell_1, \ldots, \ell_k).$$

Therefore $\mathscr{P}_{k+1}$ is a refinement of $\mathscr{P}_k$ and $\operatorname{diam}(\mathscr{P}_k(x)) \leq 2^{-k}$ for all $x \in X$.

Denote

$$\beta = \frac{1}{16n \log K} \ . \tag{35}$$

Since $K \geq 5$, we have $\beta < \frac{1}{25}$. Fix $k \in \mathbb{N}$ and $x \in X$ and observe that

$$\begin{aligned}\Pr\left[B\left(x, \frac{\beta}{2^k}\right) \subseteq \mathscr{P}_k(x)\right] &= \Pr\left[\bigcap_{\ell=1}^{k}\left\{\forall y \in B\left(x, \frac{\beta}{2^k}\right),\ j_\ell(x) = j_\ell(y)\right\}\right] \\ &\geq 1 - \sum_{\ell=1}^{k} \Pr\left[\exists y \in B\left(x, \frac{\beta}{2^k}\right),\ j_\ell(x) \neq j_\ell(y)\right]. \end{aligned} \tag{36}$$

Fix $\ell \in \{1, \ldots, k\}$. Note that

$$\begin{aligned}&\left\{\exists y \in B\left(x, \frac{\beta}{2^k}\right),\ j_\ell(x) \neq j_\ell(y)\right\} \\ &\qquad \subseteq \bigcup_{i=1}^{\infty}\bigcap_{j=1}^{i-1}\left\{r_\ell - \frac{\beta}{2^k} < d(x_i, x) \leq r_\ell + \frac{\beta}{2^k} \ \wedge\ d(x_j, x) > r_\ell + \frac{\beta}{2^k}\right\}. \end{aligned} \tag{37}$$

To prove (37), assume that there is some $y \in B\left(x, \frac{\beta}{2^k}\right)$ for which $j_\ell(x) \neq j_\ell(y)$. Let $i \in \mathbb{N}$ be the first index such that $d(x_i, x) \leq r_\ell + \frac{\beta}{2^k}$. Note that in order to prove that the event in the right hand side of (37) occurs, it suffices to show that the event

$$\bigcap_{j=1}^{i-1}\left\{r_\ell - \frac{\beta}{2^k} < d(x_i, x) \leq r_\ell + \frac{\beta}{2^k} \ \wedge\ d(x_j, x) > r_\ell + \frac{\beta}{2^k}\right\}$$

occurs, which, by the minimality of i, is equivalent to showing that $d(x_i, x) > r_\ell - \frac{\beta}{2^k}$. So, assume for the sake of contradiction that $d(x_i, x) \leq r_\ell - \frac{\beta}{2^k}$. This implies in particular that $j_\ell(x) = i$, and moreover, since $y \in B\left(x, \frac{\beta}{2^k}\right)$, we have $d(x_i, y) \leq r_\ell$, implying that $j_\ell(y) \leq i$. But, $d\left(x, x_{j_\ell(y)}\right) \leq d\left(y, x_{j_\ell(y)}\right) + d(x, y) \leq r_\ell + \frac{\beta}{2^k}$, and the minimality of i implies that $j_\ell(y) \geq i$. Thus $j_\ell(y) = i = j_\ell(x)$, contradicting our assumption on y.

Now, (37) implies that

$$\begin{aligned}&\Pr\left[\exists y \in B\left(x, \frac{\beta}{2^k}\right),\ j_\ell(x) \neq j_\ell(y)\right] \\ &\leq\ 2^{\ell+2}\int_{2^{-\ell-2}}^{e^{-\ell-1}}\left(\mu\left(B\left(x, r + \frac{\beta}{2^k}\right)\right) - \mu\left(B\left(x, r - \frac{\beta}{2^k}\right)\right)\right) \\ &\qquad \cdot\left(\sum_{i=1}^{\infty}\left(1 - \mu\left(B\left(x, r + \frac{\beta}{2^k}\right)\right)\right)^{i-1}\right) dr \\ &=\ 1 - 2^{\ell+2}\int_{\frac14 e^{-\ell b}}^{\frac12 e^{-\ell b}} \frac{\mu\left(B\left(x, r - \frac{\beta}{2^k}\right)\right)}{\mu\left(B\left(x, r + \frac{\beta}{2^k}\right)\right)} dr, \end{aligned} \tag{38}$$

Denote $h(t) \stackrel{\mathrm{def}}{=} \log \mu\left(B(x,s)\right)$. Then by Jensen's inequality we see that

$$2^{\ell+2}\int_{2^{-\ell-2}}^{2^{-\ell-1}} \frac{\mu\left(B\left(x,r-\frac{\beta}{2^k}\right)\right)}{\mu\left(B\left(x,r-\frac{\beta}{2^k}\right)\right)}dr = 2^{\ell+2}\int_{2^{-\ell-2}}^{2^{-\ell-1}} e^{h\left(r-\frac{\beta}{2^k}\right)-h\left(r+\frac{\beta}{2^k}\right)}dr$$
$$\geq \exp\left(2^{\ell+2}\int_{2^{-\ell-2}}^{2^{-\ell-1}}\left[h\left(r-\frac{\beta}{2^k}\right)-h\left(r+\frac{\beta}{2^k}\right)\right]dr\right). \quad (39)$$

The term in the exponent in (39) can be estimated as follows:

$$\int_{2^{-\ell-2}}^{2^{-\ell-1}}\left[h\left(r-\frac{\beta}{2^k}\right)-h\left(r+\frac{\beta}{2^k}\right)\right]dr = \int_{2^{-\ell-2}-\beta 2^{-k}}^{2^{-\ell-2}+\beta 2^{-k}} h(s)\,ds - \int_{2^{-\ell-1}-\beta 2^{-k}}^{2^{-\ell-1}+\beta 2^{-k}} h(s)\,ds$$
$$\geq \beta 2^{-k+1}\left[h\left(2^{-\ell-2}-\beta e^{-k}\right)-h\left(2^{-\ell-1}+\beta 2^{-k}\right)\right]. \quad (40)$$

By recalling the definition of h, a combination of (38), (39), (40) yields the bound,

$$\Pr\left[\exists y\in B\left(x,\frac{\beta}{2^k}\right),\ j_\ell(x)\neq j_\ell(y)\right]\geq 1-\left(\frac{\mu\left(B\left(x,2^{-\ell-2}-\beta 2^{-k}\right)\right)}{\mu\left(B\left(x,2^{-\ell-1}+\beta 2^{-k}\right)\right)}\right)^{\beta 2^{-(k-\ell)+3}}. \quad (41)$$

Note that since $\ell \leq k$ and $\beta \leq \frac{1}{25}$ we know that $2^{-\ell-1}+\beta 2^{-k} \leq \left(1+\frac{1}{n}\right)^{n+1}\left(2^{-\ell-2}-\beta 2^{-k}\right)$. Hence, combining the assumption (8) with (41), we see that

$$\Pr\left[\exists y\in B\left(x,\frac{\beta}{2^k}\right),\ j_\ell(x)\neq j_\ell(y)\right]\leq 1-K^{-(n+1)\beta 2^{-(k-\ell)+3}}$$
$$\leq (n+1)\beta 2^{-(k-\ell)+3}\log K \overset{(35)}{\leq} 2^{-(k-\ell)}. \quad (42)$$

Plugging (42) into (36) we see that

$$\Pr\left[B\left(x,\frac{1}{16n\log K}\cdot 2^{-k}\right)\subseteq \mathscr{P}_k(x)\right] = \Pr\left[B\left(x,\frac{\beta}{2^k}\right)\subseteq \mathscr{P}_k(x)\right]\geq 1-\sum_{\ell=1}^{k}2^{-(k-\ell)}\geq \frac{1}{2}.$$

This is precisely the statement that the partition tree $\{\mathscr{P}_k\}_{k=0}^\infty$ is $\frac{1}{16n\log K}$-padded. $\square$

The connection between the existence of padded random partition trees and the Hardy-Littlewood maximal inequality is established in the proof of Theorem 1.1.

Proof of Theorem 1.1. By a standard monotone convergence argument we may assume that R is bounded, say $R\subseteq[0,D]$ for some $D>1$. Fix $f\in L_p(X)$. By homogeneity it suffices to show that

$$\mu\left(M_R f>1\right)\lesssim C^p\left(\left(1+\frac{\log\log K}{1+\log n}\right)Q^p+K^p\right)\int_X |f|^p d\mu,$$

where $C>0$ is a universal constant and

$$Q \stackrel{\mathrm{def}}{=} \sup_{r>0}\left\|M_{R\cap[r,nr]}\right\|_{L_p(X)\to L_{p,\infty}(X)}. \quad (43)$$

By monotone convergence we may assume that f (and hence also $M_R f$) has bounded support. We would like to apply Theorem 2.1, but unfortunately there are no obvious candidates for $\mathscr{F}_k$ with which we have either (24) or (25). Nevertheless, we shall be able to proceed by replacing M_R with a slightly modified variant.

Let E be the support of f and denote

$$E' \stackrel{\mathrm{def}}{=} \{x \in X : \ d(x,E) \le D\},$$

and

$$E'' \stackrel{\mathrm{def}}{=} \{x \in X : \ d(x,E) \le 2D\}.$$

Then $E \subseteq E' \subseteq E''$ and $\mathrm{diam}(E'') \le 4D + \mathrm{diam}(E) < \infty$. Moreover the support of $M_R f$ is contained in E'. It will therefore suffice to prove that

$$\|M_R\|_{L_p(E') \to L_{p,\infty}(E'')} \lesssim \left(1 + \frac{\log\log K}{1+\log n}\right) Q + K.$$

By rescaling the metric we may assume that $\mathrm{diam}(E'') = 1$. Once this is achieved we may also assume that $R \subseteq (0,1]$, since the operator $M_{R\cap(1,\infty)}$, viewed as an operator on $L_p(E')$, is pointwise bounded by the averaging operator on E'.

Using Lemma 3.1, let $\{\mathscr{P}_k\}_{k=0}^{\infty}$ be a random partition tree on E'' which is β-padded, where

$$\beta = \frac{1}{16 n \log K}.$$

Let m be the largest integer such that $2^{-m} \le \beta$. Denote for $k \ge 0$ and $i \in \{1,2,3\}$,

$$R_k^i \stackrel{\mathrm{def}}{=} R \cap \left[2^{-(3k+i)m}, 2^{-(3k-1+i)m}\right] \quad \text{and} \quad R^i \stackrel{\mathrm{def}}{=} \bigcup_{k \in \mathbb{N}\cup\{0\}} R_k^i.$$

Thus $R = R^1 \cup R^2 \cup R^3$, which implies that

$$\mu\left(M_R f > 1\right) = \mu\left(\max\left\{M_{R^1} f, M_{R^2} f, M_{R^3} f\right\} > 1\right)$$
$$\le \mu\left(M_{R^1} f > 1\right) + \mu\left(M_{R^2} f > 1\right) + \mu\left(M_{R^3} f > 1\right). \quad (44)$$

Fix $i \in \{1,2,3\}$ and $k \in \mathbb{N}\cup\{0\}$, and define

$$E_k^i \stackrel{\mathrm{def}}{=} \left\{x \in E' : \ M_{R_k^i} f(x) > 1\right\} \setminus \bigcup_{j=0}^{k-1} \left\{x \in E' : \ M_{R_j^i} f(x) > 1\right\}.$$

Then the sets E_k^i are disjoint and

$$\mu\left(M_{R^i} f > 1\right) = \mu\left(\sup_{k \in \mathbb{N}\cup\{0\}} M_{R_k^i} f > 1\right) = \sum_{k=0}^{\infty} \mu\left(E_k^i\right). \quad (45)$$

Recalling (33), we denote

$$\widetilde{E}_k^i \stackrel{\mathrm{def}}{=} (E_k^i)_\beta^{\mathrm{pad}((3k+i+1)m)} = \left\{x \in E_k^i : \ B\left(x, \frac{\beta}{2^{(3k+i+1)m}}\right) \subseteq \mathscr{P}_{(3k+i+1)m}(x)\right\}.$$

Then by (34) we know that

$$\mathbb{E}\left[\mu\left(\widetilde{E}_k^i\right)\right] \ge \frac{\mu\left(E_k^i\right)}{2}. \quad (46)$$

Plugging (46) into (45) we see that

$$\mu\left(M_{R^i} f > 1\right) \le 2\mathbb{E}\left[\sum_{k=0}^{\infty} \mu\left(\widetilde{E}_k^i\right)\right] = 2\mathbb{E}\left[\mu\left(\sup_{k \in \mathbb{N}\cup\{0\}} \widetilde{M}_{R_k^i} f > 1\right)\right], \quad (47)$$

where $\widetilde{M}_{R_k^i}$ is the sublinear operator

$$\widetilde{M}_{R_k^i} g \stackrel{\mathrm{def}}{=} \mathbf{1}_{\widetilde{E}_k^i} M_{R_k^i} g.$$

Write $r = 2^{-(3k+i)m}$ and let $v \asymp 1 + \frac{\log\log K}{1+\log n}$ be an integer such that $2^{m/v} \le n$. By the definition of Q, for every $g \in L_p(E')$ and $t > 0$ we have

$$\mu\left(\widetilde{M}_{R_k^i} g > t\right) \le \mu\left(M_{R_k^i} g > t\right) = \mu\left(M_{R\cap[r,2^m r]} g > t\right) \le \sum_{u=0}^{v-1} \mu\left(M_{R\cap\left[r2^{\frac{um}{v}}, nr2^{\frac{um}{v}}\right]} g > t\right) \le vQ^p \frac{\|g\|_{L_p(E')}^p}{t^p}.$$

Thus,

$$g \in L_p(E') \implies \left\|\widetilde{M}_{R_k^i} g\right\|_{L_{p,\infty}(E')} \le v^{1/p} Q \|g\|_{L_p(E')}. \tag{48}$$

For every $k \in \mathbb{N}\cup\{0\}$ we let $\mathscr{F}_k \stackrel{\mathrm{def}}{=} \sigma(\mathscr{P}_k)$ be the σ-algebra generated by the partition $\mathscr{P}_k$. Then $\mathscr{F}_0 \subseteq \mathscr{F}_1 \subseteq \mathscr{F}_2 \subseteq \cdots$. We claim that for every $k \in \mathbb{N}\cup\{0\}$, if $F \in \mathscr{F}_{(3k+i+1)m}$ then

$$\mathbf{1}_F \widetilde{M}_{R_{k+1}^i}(g) = \widetilde{M}_{R_{k+1}^i}(\mathbf{1}_F g). \tag{49}$$

By the definition of $\widetilde{M}_{R_k^i}$, in order to prove (49) we have to show that for almost every $x \in E'$ we have

$$\mathbf{1}_F(x)\cdot \mathbf{1}_{\widetilde{E}_{k+1}^i}(x)\cdot M_{R_{k+1}^i}(g)(x) = \mathbf{1}_{\widetilde{E}_{k+1}^i}(x)\cdot M_{R_{k+1}^i}(\mathbf{1}_F g)(x). \tag{50}$$

It is non-trivial to check (50) only when $x \in \widetilde{E}_{k+1}^i$, in which case we are guaranteed that $B\left(x, \beta 2^{-(3k+i+1)m}\right) \subseteq \mathscr{P}_{(3k+i+1)m}(x)$. But since $F \in \mathscr{F}_{(3k+i+1)m}$, we know that $P_{(3k+i+1)m}(x)$ is either disjoint from F or contained in F. If $P_{(3k+i+1)m}(x) \subseteq F$, then for every $r \in R_{k+1}^i$,

$$B(x,r) \subseteq B\left(x, 2^{-(3k+i+2)m}\right) \subseteq B\left(x, \beta 2^{-(3k+i+1)m}\right) \subseteq \mathscr{P}_{(3k+i+1)m}(x) \subseteq F, \tag{51}$$

where we used the fact that $r \le 2^{-(3(k+1)-1+i)m}$ and $2^{-m} \le \beta$. The inclusion (51) implies that both sides of the equation (50) are equal to $M_{R_{k+1}^i}(g)(x)$. On the other hand, if $P_{(3k+i+1)m}(x)$ is disjoint from F, then $B(x,r)$ is disjoint from F for all $r \in R_{k+1}^i$, implying that both sides of the equation (50) vanish. This concludes the proof of (49).

Fix $g \in L_\infty(E')$, and extend g to a function on X whose value is 0 outside E'. Assume that

$$\left\|\mathbb{E}\left(|g|\middle|\mathscr{F}_{(3k+i+1)m}\right)\right\|_{L_\infty(E')} = 1.$$

This implies that for all $F \in \mathscr{F}_{(3k+i+1)m}$ we have

$$\int_F |g|d\mu = \int_{F\cap E'} |g|d\mu \le \mu\left(F\cap E'\right) \le \mu(F). \tag{52}$$

Fix $r \in R_k^i$ and $x \in E'$. Denote

$$F \stackrel{\mathrm{def}}{=} \bigcup\left\{C \in \mathscr{P}_{(3k+i+1)m} :\ C\cap B(x,r) \ne \emptyset\right\} \in \mathscr{F}_{(3k+i+1)m}.$$

Note that $B(x,r)\subseteq E''$, which implies that

$$F\supseteq B(x,r). \tag{53}$$

Moreover,

$$F\subseteq B\left(x, r+\sup_{C\in\mathscr{P}_{(3k+i+1)m}}\mathrm{diam}(C)\right)\subseteq B\left(x, r+2^{-(3k+i+1)m}\right)\subseteq B\left(x,\left(1+2^{-m}\right)r\right), \tag{54}$$

where in the last inclusion in (54) we used the fact that $r\in R_k^i$ implies that $r\geq 2^{-(3k+i)m}$. Hence,

$$\frac{1}{\mu(B(x,r))}\int_{B(x,r)}|g|d\mu \overset{(53)}{\leq} \frac{1}{\mu(B(x,r))}\int_{F}|g|d\mu \overset{(52)}{\leq} \frac{\mu(F)}{\mu(B(x,r))}$$
$$\overset{(54)}{\leq}\frac{\mu\left(B\left(x,\left(1+2^{-m}\right)r\right)\right)}{\mu(B(x,r))}\leq \frac{\mu\left(B\left(x,\left(1+\frac{1}{n}\right)r\right)\right)}{\mu(B(x,r))}\overset{(8)}{\leq} K, \tag{55}$$

We are now in position to apply Theorem 2.1 to the increasing sequence of σ-algebras $\left\{\mathscr{F}_{(3k+i+1)m}\right\}_{k=0}^{\infty}$ and the sublinear operators $\left\{M_{R_k^i}\right\}_{k=0}^{\infty}$, with $A=v^{1/p}Q$, due to (48), and $B=K$, due to (55):

$$\mu\left(\sup_{k\in\mathbb{N}\cup\{0\}}\widetilde{M}_{R_k^i}f>1\right)\leq\left(2^p vQ^p+2^pK^p\right)\int_X|f|^pd\mu$$
$$\lesssim\left(2^p\left(1+\frac{\log\log K}{1+\log n}\right)Q^p+2^pK^p\right)\int_X|f|^pd\mu.$$

Using (47) and (44), we therefore deduce that

$$\left[\mu\left(M_Rf>1\right)\right]^{1/p}\lesssim\left(\left(1+\frac{\log\log K}{1+\log n}\right)^{1/p}Q+K\right)\|f\|_{L_p(X)},$$

as required. □

4. An argument of E. Lindenstrauss

We now present an alternative approach to Corollary 1.2, following an argument of E. Lindenstrauss [31]. Let us first make some definitions. We fix a metric measure space (X,d,μ). Given any two radii $r,r'>0$ and a center $x\in X$, we define the enlarged ball $B(x,r,r')$ by

$$B(x,r,r')\overset{\mathrm{def}}{=}\bigcup_{y\in B(x,r)}B(y,r')=\{z\in X: d(x,y)\leq r\ \wedge\ d(y,z)\leq r' \text{ for some } y\in X\}.$$

Thus, for instance,

$$B(x,r)\subseteq B(x,r,r')\subseteq B(x,r+r'). \tag{56}$$

In analogy to [31], we say that a finite sequence of radii $0 < r_1 < r_2 < \cdots < r_k$ is *tempered* with constant $K \geq 1$ if we have the bound

$$\forall\, j \in \{1, \ldots, k\}\ \forall x \in X\ \forall y \in B(x, r_j),$$
$$\mu\left(B(x, r_j) \bigcup \left(\bigcup_{i=1}^{j-1} B\left(x, r_j, r_i\right)\right)\right) \leq K\mu\left(B\left(y, r_j\right)\right). \quad (57)$$

Theorem 4.1 (Lindenstrauss maximal inequality). *Let (X, d, μ) be a metric measure space, and let $0 < r_1 < r_2 < \ldots < r_k$ be a sequence of radii which is tempered with constant K. Then we have the weak $(1,1)$ maximal inequality*

$$\mu\left(x \in X : \max_{1 \leq j \leq k} \frac{1}{B(x, r_j)} \int_{B(x, r_j)} |f|\ d\mu > \lambda\right) \leq \frac{2e}{e-1} \frac{K}{\lambda} \|f\|_{L_1(X)}$$

for all $f \in L_1(X)$ and $\lambda > 0$.

Proof of Corollary 1.2 assuming Theorem 4.1. Assume that (X, d, μ) obeys the strong microdoubling condition (14). It is immediate to check that any sequence $0 < r_1 < r_2 < \ldots < r_k$ obeying the lacunarity condition $r_j \geq n r_{j-1}$ will be tempered with constant K, and hence by Theorem 4.1,

$$\mu\left(x \in X : \max_{1 \leq j \leq k} \frac{1}{B(x, r_j)} \int_{B(x, r_j)} |f|\ d\mu > \lambda\right) \leq \frac{2e}{e-1} \frac{K}{\lambda} \|f\|_{L_1(X)}.$$

If instead we have the lacunarity condition $r_j \geq 2r_{j-1}$, then we can sparsify this sequence into $O(\log n)$ subsequences obeying the prior lacunarity condition, and hence, by subadditivity,

$$\mu\left(x \in X : \max_{1 \leq j \leq k} \frac{1}{B(x, r_j)} \int_{B(x, r_j)} |f|\ d\mu > \lambda\right) \lesssim \frac{K \log n}{\lambda} \|f\|_{L_1(X)}.$$

From monotone convergence we then conclude (16). Similarly, any sequence obeying the lacunarity condition $r_j \geq (1 + \frac{1}{n}) r_{j-1}$ can be sparsified into $O(n \log n)$ sequences which have a lacunarity ratio of n. By monotone convergence this implies that

$$\mu\left(x \in X : \sup_{r \in (1+\frac{1}{n})^{\mathbb{Z}}} \frac{1}{B(x, r)} \int_{B(x, r)} |f|\ d\mu > \lambda\right) \lesssim \frac{K n \log n}{\lambda} \|f\|_{L_1(X)},$$

where $(1 + \frac{1}{n})^{\mathbb{Z}}$ denotes the integer powers of $1 + \frac{1}{n}$. Now note from (14) that every ball is contained in a ball whose radius is an integer power of $1 + \frac{1}{n}$, and whose measure is at most K times larger. Thus

$$Mf(x) \leq K \sup_{r \in (1+\frac{1}{n})^{\mathbb{Z}}} \frac{1}{B(x, r)} \int_{B(x, r)} |f|\ d\mu,$$

and (15) follows. □

Proof of Theorem 4.1. As in [31], this is achieved by a randomized variant of the Vitali covering argument. We may take f to be non-negative, and normalize $\lambda = 1$. For each

$j \in \{1, \ldots, k\}$, let E_j be a compact subset of X on which we have

$$x \in E_j \implies \frac{1}{B(x, r_j)} \int_{B(x,r_j)} f \, d\mu > 1. \tag{58}$$

By inner regularity it will suffice to show that

$$\mu\left(\bigcup_{j=1}^{k} E_j\right) \leq \frac{2e}{e-1} K \int_X f \, d\mu. \tag{59}$$

We establish (59) by induction on k. The case $k = 0$ is vacuously true, so suppose $k \geq 1$ and the claim has already been proven for $k-1$ (i.e, that (59) holds true for all non-negative $f \in L_1(X)$ and all sets $\{E_j\}_{j=1}^{k-1}$ satisfying (58)).

By compactness, we see that there exists an $\varepsilon > 0$ such that

$$x \in E_k \implies \mu(B(x, r_k)) > \varepsilon.$$

We then define the extended ball

$$B^*(x) \stackrel{\text{def}}{=} B(x, r_k) \bigcup \left(\bigcup_{j=1}^{k-1} B(x, r_k, r_j)\right).$$

Thus, since the sequence of radii $\{r_j\}_{j=1}^{k}$ is tempered, for all $y \in B(x, r_k)$,

$$\varepsilon < \mu\left(B^*(y)\right) \leq K\mu(B(x, r_k)). \tag{60}$$

If we then define the *intensity function*

$$p(x) \stackrel{\text{def}}{=} \inf_{y \in B(x,r_k)} \frac{1}{\mu(B^*(y))},$$

then p is a measurable function on E_k which is bounded both above and below:

$$\frac{1}{K\mu\left((B(x, r_k)\right)} \leq p(x) < \frac{1}{\varepsilon}. \tag{61}$$

We now introduce a Poisson process Σ on E_k with intensity $p(x)$. Thus Σ is a random finite subset[6] of E_k which will be almost surely finite, and more precisely, for any non-negative measurable weight $w : E_k \to \mathbb{R}_+$, the quantity $\sum_{x \in \Sigma} w(x)$ is a Poisson random variable with expectation

$$\alpha_w \stackrel{\text{def}}{=} \mathbb{E}\left[\sum_{x \in \Sigma} w(x)\right] = \int_{E_k} wp \, d\mu, \tag{62}$$

i.e., for any integer $k \geq 0$

$$\Pr\left(\sum_{x \in \Sigma} w(x) = k\right) = \frac{e^{-\alpha_w} \alpha_w^k}{k!}. \tag{63}$$

[6]If E_k contains atoms, then Σ may contain multiplicity, thus it is really a multiset rather than a set in this case. One way to create Σ is to let N be a Poisson random variable with expectation $P \stackrel{\text{def}}{=} \int_{E_k} p d\mu$ and then let $\Sigma = \{x_1, \ldots, x_N\}$ where $x_1, \ldots, x_N$ are iid elements of E chosen using the probability distribution $p d\mu_Y / P$.

Now we define the random sets

$$E' \stackrel{\text{def}}{=} \bigcup_{x\in\Sigma} B^*(x) \quad \text{and} \quad F \stackrel{\text{def}}{=} \bigcup_{x\in\Sigma} B(x, r_k).$$

Then,

$$\mu\left(\bigcup_{j=1}^{k} E_j\right) \le \mu(E_k) + \mu(E') + \mu\left(\bigcup_{j=1}^{k-1} E_j \setminus E'\right). \tag{64}$$

Let us investigate the third term in (64). Fix $j \in \{1, \ldots, k-1\}$. If $x \in E_j \setminus E'$, then

$$\frac{1}{B(x, r_j)} \int_{B(x,r_j)} f \, d\mu > 1.$$

But, since $x \notin E'$ it follows from our definitions that $B(x, r_j)$ is disjoint from F. Thus we have

$$\frac{1}{B(x, r_j)} \int_{B(x,r_j)} f \mathbf{1}_{X\setminus F} \, d\mu > 1.$$

We can therefore apply the induction hypothesis to the sets $\{E_j \setminus E'\}_{j=1}^{k-1}$ and the function $f\mathbf{1}_{X\setminus F}$, and conclude that

$$\mu\left(\bigcup_{j=1}^{k-1} E_j \setminus E'\right) \le \frac{2e}{e-1} K \int_{X\setminus F} f \, d\mu.$$

It follows from (64) that it suffices to show that

$$\mu(E_k) + \mathbb{E}\left[\mu(E')\right] \le \mathbb{E}\left[\mu(E_k) + \mu(E')\right] \le \frac{2e}{e-1} K \mathbb{E}\left[\int_F f \, d\mu\right]. \tag{65}$$

Now, applying (62) and (63) with $w \stackrel{\text{def}}{=} 1/p$, we have

$$\mu(E_k) = \mathbb{E}\left[\sum_{x\in\Sigma} \frac{1}{p(x)}\right],$$

while from definition of E' we have

$$\mu(E') \le \sum_{x\in\Sigma} \mu\left(B^*(x)\right) = \sum_{x\in\Sigma} \frac{1}{p(x)}.$$

Thus, in order to prove (65) it suffices to show that

$$\mathbb{E}\left[\sum_{x\in\Sigma} \frac{1}{p(x)}\right] \le \frac{e}{e-1} K \mathbb{E}\left[\int_F f \, d\mu\right]. \tag{66}$$

¿From (58) we know that for all $x \in \Sigma$,

$$\frac{1}{p(x)} < \frac{1}{p(x)\mu(B(x, r_k))} \int_X \mathbf{1}_{B(x,r_k)} f \, d\mu,$$

and hence

$$\mathbb{E}\left[\sum_{x\in\Sigma} \frac{1}{p(x)}\right] \le \int_X \left(\mathbb{E}\left[\sum_{x\in\Sigma} \frac{1}{p(x)\mu(B(x, r_k))} \mathbf{1}_{B(x,r_k)}\right]\right) f \, d\mu. \tag{67}$$

Fix $y \in X$. From (62) with $w(x) = \frac{\mathbf{1}_{B(x,r_k)}(y)}{p(x)\mu(B(x,r_k))}$, we see that

$$\mathbb{E}\left[\sum_{x\in\Sigma} \frac{1}{p(x)\mu(B(x,r_k))}\mathbf{1}_{B(x,r_k)}(y)\right] = \int_{E_k\cap B(y,r_k)} \frac{1}{\mu(B(x,r_k))}\, d\mu(x). \tag{68}$$

By substituting (68) into (67), we see that in order to prove (66) it will suffice to prove the pointwise estimate

$$\int_{E_k\cap B(y,r_k)} \frac{1}{\mu(B(x,r_k))}\, d\mu(x) \le \frac{eK}{e-1}\mathbb{E}\left[\mathbf{1}_F(y)\right], \tag{69}$$

for all $y \in X$.

Now observe that the definition of F implies that $\mathbf{1}_F(y) = 1$ if and only if $|\Sigma\cap B(y,r_k)| \ge 1$. But, recall from (62) (using $w(x) = \mathbf{1}_{B(y,r_k)}(x)$) that $|\Sigma \cap B(y,r_k)|$ is a Poisson random variable with expectation

$$\alpha(y) \stackrel{\mathrm{def}}{=} \int_{E_k} \mathbf{1}_{B(y,r_k)}p\, d\mu = \int_{E_k\cap B(y,r_k)} p(x)\, d\mu(x), \tag{70}$$

and thus

$$\mathbb{E}\left[\mathbf{1}_F(y)\right] = 1 - e^{-\alpha(y)}. \tag{71}$$

A combination of (61) and (70) yields the bound

$$\int_{E_k\cap B(y,r_k)} \frac{1}{\mu(B(x,r_k))}\, d\mu(x) \le K\alpha(y). \tag{72}$$

The definition of $p(x)$ implies that if $y \in B(x,r_k)$ then $p(x) \le \frac{1}{\mu(B^*(y))} \le \frac{1}{\mu(B(y,r_k))}$, since $B^*(y) \supseteq B(x,r_k)$. In combination with (70), we deduce that $\alpha(y) \le 1$. But, the function $\alpha \mapsto \frac{1-e^{-\alpha}}{\alpha}$ is decreasing on $[0,\infty)$, and therefore $1 - e^{-\alpha(y)} \ge (1-e^{-1})\alpha(y)$. This, in combination with (71) and (72), implies (69), and completes the proof of Theorem 4.1. □

As observed in [31], the above argument allows us to extract a good maximal inequality for sufficiently sparse subsequences of radii if the situation is sufficiently "amenable". In our current context, the analogue for amenability is in fact subexponential growth:

Corollary 4.2. *Let (X,d,μ) be a metric measure space such that $\mu(B(x,r)$ is independent of $x \in X$ for all $r > 0$. Suppose also that we have the sub-exponential growth condition*

$$\lim_{r\to\infty} \frac{\log \mu(B(x,r))}{r} = 0 \tag{73}$$

for any $x \in X$ (note that our assumption implies that the choice of x is in fact irrelevant). Then there exists a sequence of radii $0 < r_1 < r_2 < \dots$ tending to infinity such that we have the maximal inequality

$$f \in L_1(X) \implies \left\|\sup_{k\ge1} A_{r_k}|f|\right\|_{L_{1,\infty}(X)} \le 4\|f\|_{L_1(X)},$$

where the averaging operators A_r are given by $A_r g \stackrel{\mathrm{def}}{=} \frac{1}{B(x,r)}\int_{B(x,r)} |g|\, d\mu$.

Proof. We construct the radii recursively as follows. We set $r_1 \stackrel{\text{def}}{=} 1$. If $r_1, \ldots, r_k$ have already been chosen, we choose $r_{k+1} > \max\{r_k, k\}$ so that

$$\log \mu\left(B\left(x, r_{k+1} + r_k\right)\right) \leq \mu\left(B\left(x, r_{k+1}\right)\right) + 0.001$$

for any $x \in X$. Such a radius must exist, since otherwise one would easily contradict (73). The sequence of radii is tempered with constant $K = e^{0.001}$, and the claim follows since $\frac{2K}{1-e^{-1}} < 4$. □

5. The infinite tree

Fix $k \geq 2$ and let T be the infinite rooted k-ary tree with the usual graph metric and the counting measure μ. In this section we prove Theorem 1.5. The first (standard) step is to replace the Hardy-Littlewood maximal function with the spherical maximal function

$$M^\circ f(x) \stackrel{\text{def}}{=} \sup_{r \geq 0} \frac{1}{|S(x,r)|} \sum_{y \in S(x,r)} |f(y)|,$$

where $S(x,r)$ is the sphere

$$S(x,r) \stackrel{\text{def}}{=} \{y \in T : d(x,y) = r\}.$$

Since every ball can be written as the disjoint union of spheres, we have the pointwise estimate

$$Mf(x) \leq M^\circ f(x),$$

and so it suffices to show that

$$\mu\left(x \in T : M^\circ f(x) \geq \lambda\right) \lesssim \frac{1}{\lambda} \|f\|_{L_1(T)}, \tag{74}$$

for all $f \in L_1(T)$ and $\lambda > 0$.

Our arguments rely on the following expander-type estimate. We use $|E| = \mu(E)$ to denote the cardinality of a finite set $E \subseteq T$.

Lemma 5.1. *Let E, F be finite subsets of T and let $r \geq 0$ be an integer. Then*

$$|\{(x,y) \in E \times F : \ d(x,y) = r\}| \leq 2|E|^{1/2}|F|^{1/2}k^{r/2}.$$

This bound should be compared against the "trivial" bounds of $|E|k^r$ and $|F|k^r$. It is superior when $|E|/|F|$ lies between k^r and k^{-r}. By setting E and F equal to concentric spheres one can verify that the bound is essentially sharp in this case.

Proof. Let us subdivide $T = \bigcup_{j=0}^\infty T_j$, where T_j is the generation of the tree at depth j (thus for instance $|T_j| = k^j$). We then define $E_j \stackrel{\text{def}}{=} E \cap T_j$ and $F_j \stackrel{\text{def}}{=} F \cap T_j$. Observe that in order for an element in E_j and an element in F_i to have distance exactly r, we must have $i = j + r - 2m$ for some $m \in \{0, \ldots, r\}$. Thus we can write

$$|\{(x,y) \in E \times F : d(x,y) = r\}| = \sum_{m=0}^{r} \sum_{\substack{i,j \in \mathbb{N} \cup \{0\} \\ i = j + r - 2m}} \left|\{(x,y) \in E_j \times F_i : d(x,y) = r\}\right|. \tag{75}$$

Fix $m \in \{0, \ldots, r\}$ and $i, j \in \mathbb{N} \cup \{0\}$ such that $i = j + r - 2m$. Observe that if $x \in T_j$ and $y \in T_i$ are at distance r in T, then the m^{th} parent of x equals the $(r-m)^{th}$ parent of y. From this we conclude that for each $x \in T_j$ there are at most k^{r-m} elements of $y \in T_i$ with

$d(x,y)=r$, and conversely for each $y\in T_i$ there are at most k^m elements of $x\in T_j$ with $d(x,y)=r$. Thus

$$|\{(x,y)\in E_j\times F_i : d(x,y)=r\}|\le \min\left\{k^{r-m}|E_j|, k^m|F_i|\right\}. \tag{76}$$

A combination of (75) and (76) implies that our task is therefore to show that

$$\sum_{m=0}^{r}\sum_{\substack{i,j\in\mathbb{N}\cup\{0\}\\ i=j+r-2m}}\min\left\{k^{r-m}|E_j|, k^m|F_i|\right\}\le 2|E|^{1/2}|F|^{1/2}k^{r/2}. \tag{77}$$

If we write $c_j\stackrel{\mathrm{def}}{=}\frac{|E_j|}{k^j}$ and $d_j\stackrel{\mathrm{def}}{=}\frac{|F_j|}{k^j}$ for $j\ge 0$ and $c_j\stackrel{\mathrm{def}}{=}d_j\stackrel{\mathrm{def}}{=}0$ for $j<0$ then we have

$$\sum_{j=0}^{\infty}k^jc_j=|E|\quad\text{and}\quad\sum_{j=0}^{\infty}k^jd_j=|F|, \tag{78}$$

and we have

$$\begin{aligned}\sum_{m=0}^{r}\sum_{\substack{i,j\in\mathbb{N}\cup\{0\}\\ i=j+r-2m}}\min\left\{k^{r-m}|E_j|, k^m|F_i|\right\}&=\sum_{m=0}^{r}\sum_{\substack{i,j\in\mathbb{N}\cup\{0\}\\ i=j+r-2m}}k^{(i+j+r)/2}\min\{c_j,d_i\}\\ &\le k^{r/2}\sum_{i,j=0}^{\infty}k^{(i+j)/2}\min\{c_j,d_i\}. \end{aligned}\tag{79}$$

A combination of (78) and (79) shows that in order to prove (77) it will suffice to show that

$$\sum_{i,j=0}^{\infty}k^{(i+j)/2}\min\{c_j,d_i\}\le 2\left(\sum_{j\ge 0}k^jc_j\right)^{1/2}\left(\sum_{i\ge 0}k^id_i\right)^{1/2}.$$

To prove this inequality, let α be a real parameter to be chosen later, and estimate

$$\sum_{i,j=0}^{\infty}k^{(i+j)/2}\min\{c_j,d_i\}\le\sum_{\substack{i,j\in\mathbb{N}\cup\{0\}\\ i<j+\alpha}}k^{(i+j)/2}c_j+\sum_{\substack{i,j\in\mathbb{N}\cup\{0\}\\ i\ge j+\alpha}}k^{(i+j)/2}d_i\le\sum_{j=0}^{\infty}k^{j+\frac{\alpha}{2}}c_j+\sum_{i=0}^{\infty}k^{i-\frac{\alpha}{2}}d_i.$$

Optimising in α we obtain the required result. □

For each $r\ge 0$, let A_r° denote the spherical averaging operator

$$A_r^\circ f(x)\stackrel{\mathrm{def}}{=}\frac{1}{\mu(S(x,r))}\sum_{y\in S(x,r)}|f(y)|.$$

Thus $M^\circ f(x)=\sup_{r\ge 0}A_r^\circ f(x)$. We can use Lemma 5.1 to obtain a distributional estimate on A_r°.

Lemma 5.2. *Let $f\in L_1(T)$, $r>0$ and $\lambda>0$. Then*

$$\mu\left(A_r^\circ f\ge\lambda\right)\lesssim\sum_{\substack{n\in\mathbb{N}\cup\{0\}\\ 1\le 2^n\le 2k^r}}\sqrt{\frac{2^n}{k^r}}\cdot 2^n\mu\left(|f|\ge 2^{n-1}\lambda\right).$$

Proof. We may take f to be non-negative. By dividing f by λ we may normalize $\lambda = 1$. We bound

$$f \leq \frac{1}{2} + \sum_{\substack{n\in\mathbb{N}\cup\{0\} \\ 1\leq 2^n \leq k^r}} 2^n \mathbf{1}_{E_n} + f\mathbf{1}_{\{f\geq \frac{1}{2}k^r\}}, \tag{80}$$

where E_n is the sublevel set

$$E_n \stackrel{\text{def}}{=} \left\{2^{n-1} \leq f < 2^n\right\}. \tag{81}$$

Hence

$$A_r^\circ f \leq \frac{1}{2} + \sum_{\substack{n\in\mathbb{N}\cup\{0\} \\ 1\leq 2^n \leq k^r}} 2^n A_r^\circ \left(\mathbf{1}_{E_n}\right) + A_r^\circ \left(f\mathbf{1}_{\{f\geq \frac{1}{2}k^r\}}\right). \tag{82}$$

Since $\mu(S(x,r)) \leq k^r$ we see that

$$\mu\left(A_r^\circ \left(f\mathbf{1}_{\{f\geq \frac{1}{2}k^r\}}\right) \neq 0\right) \leq k^r \mu\left(f \geq \frac{1}{2}k^r\right). \tag{83}$$

Thus we have

$$\mu\left(A_r^\circ f \geq 1\right) \stackrel{(82)\wedge(83)}{\leq} \mu\left(\sum_{\substack{n\in\mathbb{N}\cup\{0\} \\ 1\leq 2^n \leq k^r}} 2^n A_r^\circ \left(\mathbf{1}_{E_n}\right) \geq \frac{1}{2}\right) + k^r \mu\left(f \geq \frac{1}{2}k^r\right).$$

Note that if

$$\sum_{\substack{n\in\mathbb{N}\cup\{0\} \\ 1\leq 2^n \leq k^r}} 2^n A_r^\circ \left(\mathbf{1}_{E_n}\right) \geq \frac{1}{2}$$

then we necessarily have for some $n \in \mathbb{N}$ such that $1 \leq 2^n \leq k^r$,

$$A_r^\circ \left(\mathbf{1}_{E_n}\right) \geq \frac{1}{2^{n+4}} \left(\frac{2^n}{k^r}\right)^{1/4}.$$

Indeed, otherwise we have

$$\frac{1}{2} \leq \sum_{\substack{n\in\mathbb{N}\cup\{0\} \\ 1\leq 2^n \leq k^r}} 2^n A_r^\circ \left(\mathbf{1}_{E_n}\right) \leq \frac{1}{16} \sum_{\substack{n\in\mathbb{N}\cup\{0\} \\ 1\leq 2^n \leq k^r}} \left(\frac{2^n}{k^r}\right)^{1/4} \leq \frac{2^{1/4}k^{r/4} - 1}{16k^{r/4}\left(2^{1/4} - 1\right)} < \frac{1}{2},$$

which is a contraction. Thus

$$\mu\left(A_r^\circ f \geq 1\right) \leq \sum_{\substack{n\in\mathbb{N}\cup\{0\} \\ 1\leq 2^n \leq k^r}} \mu(F_n) + k^r \mu\left(f \geq \frac{1}{2}k^r\right), \tag{84}$$

where

$$F_n \stackrel{\text{def}}{=} \left\{A_r^\circ \left(\mathbf{1}_{E_n}\right) \geq \frac{1}{2^{n+4}} \left(\frac{2^n}{k^r}\right)^{1/4}\right\}.$$

Note that F_n is finite and observe that

$$\frac{1}{k^r} \left|\{(x,y) \in E_n \times F_n : d(x,y) = r\}\right| = \sum_{y\in F_n} A_r^\circ \left(\mathbf{1}_{E_n}\right)(y) \geq \frac{\mu(F_n)}{2^{n+4}} \left(\frac{2^n}{k^r}\right)^{1/4}.$$

Applying Lemma 5.1 we conclude that

$$\frac{\mu(F_n)}{2^{n+4}}\left(\frac{2^n}{k^r}\right)^{1/4}\le 2\sqrt{\frac{\mu(E_n)\mu(F_n)}{k^r}}.$$

Hence

$$\mu(F_n)\le 2^{10}\sqrt{\frac{2^n}{k^r}}\cdot 2^n\mu(E_n).$$

Plugging this estimate into (84), we obtain the required result. □

Proof of Theorem 1.5. Now we prove (74). Since $M^\circ f=\sup_{r\ge 0}A_r^\circ f$, Lemma 5.2 implies that

$$\mu\left(M^\circ f\ge\lambda\right)\le\sum_{r=0}^{\infty}\mu\left(A_r^\circ f\ge\lambda\right)\lesssim\sum_{r=0}^{\infty}\sum_{\substack{n\in\mathbb{N}\cup\{0\}\\ 1\le 2^n\le 2k^r}}\sqrt{\frac{2^n}{k^r}}\cdot 2^n\mu\left(|f|\ge 2^{n-1}\lambda\right)$$

$$=\sum_{x\in T}\sum_{n=0}^{\infty}\left(\sum_{\substack{r\in\mathbb{N}\cup\{0\}\\ k^r\ge 2^{n-1}}}\frac{1}{k^{r/2}}\right)2^{3n/2}\mathbf{1}_{\{|f(x)|\ge 2^{n-1}\lambda\}}\lesssim\sum_{x\in T}\sum_{n=0}^{\infty}2^n\mathbf{1}_{\{|f(x)|\ge 2^{n-1}\lambda\}}\lesssim\sum_{x\in T}\frac{1}{\lambda}|f(x)|,$$

which is (74), as desired. The proof of Theorem 1.5 is complete. □

6. Sharpness

The purpose of this section is to prove Theorem 1.3 and Theorem 1.4.

6.1. A preliminary construction.
Before we exhibit the full examples, we first need a preliminary example of a maximal operator associated to a finite Abelian group (but not to a metric) which has bad weak $(1,1)$ behavior.

Proposition 6.2 (Preliminary example)**.** *Let q be a power of an odd prime, and let $\mathbb{F}_q$ be the finite field with q elements. If q is sufficiently large then there exists a vector space X_q over $\mathbb{F}_q$ with counting measure μ and dimension $m=\dim(X_q)\le\sqrt{q}$, and disjoint sets $\{E_z\subseteq X_q\}_{z\in\mathbb{F}_q}$ which are symmetric around the origin (i.e. $x\in E_z$ if and only if $-x\in E_z$) with measure*

$$z\in F_q\implies\frac{1}{2q}\mu(X_q)<\mu(E_z)<\frac{2}{q}\mu(X_q),\tag{85}$$

and such that the maximal function

$$M_qf(x)\stackrel{\mathrm{def}}{=}\max_{z\in\mathbb{F}_q}\frac{1}{\mu(E_z)}\int_{E_z}|f(x+y)|d\mu(y)$$

obeys the bounds

$$\|M_qf\|_{L_p(X_q)}\lesssim\left(\frac{p}{p-1}\right)^2\|f\|_{L_p(X_q)}\tag{86}$$

for all $1<p\le\infty$, but such that

$$\|M_q\|_{L_1(X_q)\to L_{1,\infty}(X_q)}>\frac{q}{2}.\tag{87}$$

Furthermore, there exists a one-dimensional subspace W_{-m+1} in X_q with the property that for all $z \in \mathbb{F}_q$

$$\mu(W_{-m+1} + E_z) \geq \frac{1}{4}\mu(X_q), \tag{88}$$

where $W_{-m+1} + E_z$ is the Minkowski sum of W_{-m+1} and E_z.

Remark 6.1. The dimension bound $\dim(X_q) \leq \sqrt{q}$ is not necessary for Theorem 1.3, but will be useful for proving Theorem 1.4. Conversely, the property (88) is used for Theorem 1.3 but not for Theorem 1.4. Even though our choice of notation for W_{-m+1} seems somewhat cumbersome at this juncture, it will become convenient when we apply Proposition 6.2 in Section 6.3.

Proof of Proposition 6.2. Let m be the largest integer less than $\sqrt{q}$. We set $X_q \stackrel{\mathrm{def}}{=} \mathbb{F}_q^m$ to be the m-dimensional vector space over $\mathbb{F}_q$, with counting measure μ. On this space we consider the non-degenerate quadratic form[7] $Q : X_q \to \mathbb{F}_q$ by

$$Q(x_1, \ldots, x_m) \stackrel{\mathrm{def}}{=} x_1^2 + \ldots + x_m^2.$$

Define for $z \in \mathbb{F}_q$

$$E_z \stackrel{\mathrm{def}}{=} \{x \in \mathbb{F}_q^m : Q(x) = z\} = Q^{-1}(z).$$

Clearly E_z is symmetric around the origin.

Let $\mathbb{F}_q^*$ denote the dual of the additive group of $\mathbb{F}_q$. Fix a non-trivial character $\chi \in \mathbb{F}_q^* \setminus \{1\}$. Then a standard Gauss sum argument (see Lemma 4.14 in [51]) shows that since q is odd,

$$\left| \sum_{x \in \mathbb{F}_q} \chi\left(yx^2\right) \right| = \sqrt{q} \tag{89}$$

for every $y \in \mathbb{F}_q \setminus \{0\}$.

For every $x = (x_1, \ldots, x_m), x' = (x_1', \ldots, x_m') \in X_q$ write $\langle x, x' \rangle \stackrel{\mathrm{def}}{=} \sum_{j=1}^m x_j x_j' \in \mathbb{F}_q$. Then for every $\eta \in X_q$ and $y \in \mathbb{F}_q \setminus \{0\}$ we have (using the fact that q is odd),

$$\begin{aligned}\left| \int_{X_q} \chi\left(yQ(x) + \langle \eta, x \rangle\right) d\mu(x) \right| &= \prod_{j=1}^m \left| \sum_{x_j \in \mathbb{F}_q} \chi\left(yx_j^2 + \eta_j x_j\right) \right| \\ &= \prod_{j=1}^m \left| \sum_{x_j \in \mathbb{F}_q} \chi\left(y\left(x_j + \frac{\eta_j}{2y}\right)^2 \right) \right| \stackrel{(89)}{=} q^{m/2} = \frac{\mu(X_q)}{q^{m/2}}.\end{aligned} \tag{90}$$

Consider the elementary identity

$$\mathbf{1}_{E_z}(x) = \frac{1}{q} \sum_{y \in \mathbb{F}_q} \chi(-yz)\chi\left(yQ(x)\right). \tag{91}$$

For every $\eta \in X_q$ and $z \in \mathbb{F}_q$ write

$$\widehat{\mathbf{1}}_{E_z}(\eta) \stackrel{\mathrm{def}}{=} \frac{1}{\mu(X_q)} \int_{X_q} \mathbf{1}_{E_z}(x)\chi\left(\langle \eta, x \rangle\right) d\mu(x).$$

[7] One could also use here a random symmetric function from F_q^m to F_q if desired; the key features of Q that we shall need are that it is even, and its Fourier coefficients are all small.

Then

$$\left|\frac{\mu\left(E_z\right)}{\mu(X_q)}-\frac{1}{q}\right| = \left|\widehat{\mathbf{1}}_{E_z}(0)-\frac{1}{q}\right| \overset{(91)}{\leq} \frac{1}{q\mu(X_q)}\sum_{y\in\mathbb{F}_q\setminus\{0\}}\left|\int_{X_q}\chi\left(yQ(x)+\langle\eta,x\rangle-yz\right)d\mu(x)\right|$$
$$\overset{(90)}{\leq} \frac{1}{q^{m/2}} \leq \frac{\sqrt{q}}{q^{\frac{1}{2}\sqrt{q}}} \leq \frac{1}{2q}, \quad (92)$$

provided that q is large enough. This proves (85). Moreover, for every $\eta \in X_q \setminus \{0\}$,

$$\left|\widehat{\mathbf{1}}_{E_z}(\eta)\right| \overset{(91)}{\leq} \frac{1}{q\mu(X_q)}\sum_{y\in\mathbb{F}_q\setminus\{0\}}\left|\int_{X_q}\chi\left(yQ(x)+\langle\eta,x\rangle\right)d\mu(x)\right| \overset{(90)}{\leq} \frac{1}{q^{m/2}} \leq \frac{\sqrt{q}}{q^{\frac{1}{2}\sqrt{q}}}. \quad (93)$$

Consider the averaging operator

$$A_zf(x) \overset{\text{def}}{=} \frac{1}{\mu(E_z)}\int_{E_z} f(x+y)d\mu(y).$$

Inequalities (92) and (93), combined with Parseval's identity, imply the L_2 bound

$$\|f-A_zf\|_{L_2(X_q)} \leq \frac{\mu(X_q)}{\mu(E_z)}\cdot\max_{\eta\in X_q\setminus\{0\}}\left|\widehat{\mathbf{1}}_{E_z}(\eta)\right|\cdot\|f\|_{L_2(X_q)} \leq \frac{2\|f\|_{L_2(X_q)}}{q^{\frac{m}{2}-1}}. \quad (94)$$

On the other hand, since A_z is a contraction in L_1, we have

$$\|f-A_zf\|_{L_1(X_q)} \leq 2\|f\|_{L_1(X_q)}. \quad (95)$$

Interpolating between (94) and (95) (see [57]) we get that for every $1 \leq p \leq 2$,

$$\|f-A_zf\|_{L_p(X_q)} \leq 2^{\frac{2}{p}-1}\cdot\left(\frac{2}{q^{\frac{m}{2}-1}}\right)^{2-\frac{2}{p}}\|f\|_{L_p(X_q)} = 2\left(q^{1-\frac{m}{2}}\right)^{2-\frac{2}{p}}\|f\|_{L_p(X_q)}. \quad (96)$$

Hence

$$\left\|\max_{z\in\mathbb{F}_q}\left||f|-A_z\left(|f|\right)\right|\right\|_{L_p(X_q)} \leq \left(\sum_{z\in\mathbb{F}_q}\left\|\left||f|-A_z\left(|f|\right)\right|\right\|_{L_p(X_q)}^p\right)^{1/p}$$
$$\overset{(96)}{\leq} 2q^{1/p}\left(q^{1-\frac{m}{2}}\right)^{2-\frac{2}{p}}\|f\|_{L_p(X_q)}.$$

Thus

$$\|M_qf\|_{L_p(X_q)} \leq \left(1+2q^{1/p}\left(q^{1-\frac{m}{2}}\right)^{2-\frac{2}{p}}\right)\|f\|_{L_p(X_q)}$$
$$\leq \left(1+2q^{1/p}\left(q^{\frac{3}{2}-\frac{\sqrt{q}}{2}}\right)^{2-\frac{2}{p}}\right)\|f\|_{L_p(X_q)} \lesssim \frac{\|f\|_{L_p(X_q)}}{(p-1)^2}. \quad (97)$$

The last step in (97) can be proved as follows: for $q \geq 36$, the term $1+2q^{1/p}\left(q^{\frac{3}{2}-\frac{\sqrt{q}}{2}}\right)^{2-\frac{2}{p}}$ is $\lesssim 1+q^{1-\varepsilon-\frac{\varepsilon\sqrt{q}}{2}}$, where we write $\frac{1}{p}=1-\varepsilon$. Now consider the cases $\varepsilon \geq \frac{2}{\sqrt{q}}$ and $\varepsilon < \frac{2}{\sqrt{q}}$ separately. The bound (97) proves (86) when $1 \leq p \leq 2$. The case $p > 2$ follows from a similar interpolation argument, using trivial bound $\|M_qf\|_{L_\infty(X_q)} \leq \|f\|_{L_\infty(X_q)}$.

To prove (87), let $f \stackrel{\mathrm{def}}{=} \mathbf{1}_{\{0\}}$ be the indicator function of the origin 0. Then $\|f\|_{L_1(X)} = 1$. Since the sets $\{E_z\}_{z\in\mathbb{F}_q}$ cover X_q, we see from (85) that $M_q f(x) > \frac{q}{2\mu(X_q)}$ for all $x \in X_q$, and (87) follows (setting λ slightly larger than $\frac{q}{2\mu(X_q)}$).

Finally, let W_{-m+1} be the span of the first basis vector $e_1 \in X_q = \mathbb{F}_q^m$. Let S denote the set of squares in $\mathbb{F}_q$, i.e. $S \stackrel{\mathrm{def}}{=} \{x^2 : x \in \mathbb{F}_q\}$. Since q is odd, $|S| = \frac{q+1}{2}$. Observe that $(x_1, \ldots, x_m)$ lies in $E_z + W_{-m+1}$ if and only if $x_2^2 + \ldots + x_m^2$ is in $z - S$. Arguing as in (92) we deduce that

$$\begin{aligned}
\mu\left(E_z + W_{-m+1}\right) &= \sum_{s\in S} \mu\left((x_1, x_2, \ldots, x_m) \in \mathbb{F}_q^m :\ x_2^2 + \cdots + x_m^2 = z - s\right)\\
&= \sum_{s\in S} q\left|\{(x_2, \ldots, x_m) \in \mathbb{F}_q^{m-1} :\ x_2^2 + \cdots + x_m^2 = z - s\}\right|\\
&\geq \sum_{s\in S} q\left(\frac{|\mathbb{F}_q^{m-1}|}{q} - \frac{|\mathbb{F}_q^{m-1}|}{q^{(m-1)/2}}\right)\\
&= \mu(X_q)\left(\frac{q+1}{2q} - \frac{q+1}{2q^{(m-1)/2}}\right).
\end{aligned}$$

This establishes the bound (88) for q sufficiently large. $\square$

Remark 6.2. This example once again demonstrates the (well-known) fact that L_2-type smoothing estimates, such as those arising from smallness of Fourier coefficients, can imply L_p maximal bounds by standard interpolation arguments, but do not necessarily imply weak-type $(1,1)$ bounds.

6.3. **The doubling example.** We now prove Theorem 1.3. The claim is trivial for $K \leq 48$, so we will assume $K \geq 48$. By Bertrand's postulate we may find an odd prime q between $K/4$ and $K/2$, which we now fix. We then let $\mathbb{F}_q$, X_q and $\{E_z\}_{z\in\mathbb{F}_q}$ be as in Proposition 6.2. Fix an arbitrary enumeration of the points in $\mathbb{F}_q$, say $\mathbb{F}_q = \{z_1, \ldots, z_q\}$ and write $E_{z_j} = E_j$ (this will not create any ambiguity in what follows). It will also be convenient to set $E_0 = \{0\}$. The maximal function M_q in Proposition 6.2 is not associated to a metric, let alone one with the doubling property (9), since the sets E_j are not nested. However, this can be remedied by extending the space X_q in the following fashion.

We let $X \stackrel{\mathrm{def}}{=} X_q \times \mathbb{F}_q^q$ be the Cartesian product of X_q with the vector space $\mathbb{F}_q^q$, with counting measure μ. We also let

$$\{0\} = V_0 \subseteq V_1 \subseteq \ldots \subseteq V_q = \mathbb{F}_q^q$$

be the standard flag in $\mathbb{F}_q^q$, thus V_j is the span of $\{e_1, \ldots, e_j\}$ for $j \in \{0, \ldots, q\}$, where $e_1, \ldots, e_q$ is the standard basis of $\mathbb{F}_q^q$. In particular

$$j \in \{0, \ldots, q\} \implies \mu(V_j) = q^j. \tag{98}$$

Recall that X_q is itself a vector space $\mathbb{F}_q^m$ over $\mathbb{F}_q$, thus we have another flag

$$\{0\} = W_{-m} \subseteq W_{-m+1} \subseteq \ldots \subseteq W_0 = X_q,$$

where

$$j \in \{0, \ldots, m\} \implies \mu(W_{-j}) = \frac{\mu(X_q)}{q^j}. \tag{99}$$

We can ensure that W_{-m+1} is the one-dimensional subspace mentioned in Proposition 6.2.

For $v \in \mathbb{F}_q$ let $j(v)$ denote the minimal $j \in \{0,\ldots,q\}$ such that $v \in V_j$. For $u \in X_q$ and $j \in \{0,\ldots,q\}$, let $\ell_j(u)$ be the maximal $\ell \in \{0,\ldots,m\}$ such that $u \in E_j + W_{-\ell}$. Now, for $(u,v),(u',v') \in X$ define

$$d((u,v),(u',v')) \stackrel{\text{def}}{=} 4^{j(v-v')}\mathbf{1}_{\{v\neq v'\}} + 2^{-\ell_{j(v-v')}(u-u')}\mathbf{1}_{\{u\neq u'\}}. \tag{100}$$

We claim that d is a translation invariant metric on X. The translation invariance and non-degeneracy of d are immediate from the definition. The symmetry of d follows from the fact that the $E_j \subseteq X_q$ are symmetric around the origin. It therefore remains to verify that for all $x,y \in X$ we have $d(x+y,0) \le d(x,0)+d(y,0)$. Write $x=(u,v)$, $y=(u',v')$, $j=j(v)$, $j'=j(v')$, $\ell=\ell_j(u)$, $\ell'=\ell_{j'}(u')$. Without loss of generality $j \ge j'$. Then $v \in V_j$ and $v' \in V_{j'} \subseteq V_j$. So, $v+v' \in V_j$, i.e., $j(v+v') \le j$. Denoting $\ell''=\ell_{j(v+v')}(u+u')$, we see that it suffices to prove the inequality

$$4^j\mathbf{1}_{\{v+v'\neq 0\}} + 2^{-\ell''}\mathbf{1}_{\{u+u'\neq 0\}} \le 4^j\mathbf{1}_{\{v\neq 0\}} + 2^{-\ell}\mathbf{1}_{\{u\neq 0\}} + 4^{j'}\mathbf{1}_{\{v'\neq 0\}} + 2^{-\ell'}\mathbf{1}_{\{u'\neq 0\}} \tag{101}$$

If $j' \ge 1$ then $v,v' \neq 0$, and (101) holds since $4^{j'} \ge 4 \ge 2^{-\ell''}$. On the other hand, if $j'=0$ (equivalently $v'=0$) then by definition $u' \in W_{-\ell'}$. Since $u \in E_j + W_{-\ell}$, it follows that $u+u' \in E_j + W_{-\ell} + W_{-\ell'} = E_{j(v+v')} + W_{-\min\{\ell,\ell'\}}$. Thus $\ell'' \ge \min\{\ell,\ell'\}$, and (101) follows from the trivial inequality $2^{-\min\{\ell,\ell'\}}\mathbf{1}_{\{u+u'\neq 0\}} \le 2^{-\ell}\mathbf{1}_{\{u\neq 0\}} + 2^{-\ell'}\mathbf{1}_{\{u'\neq 0\}}$.

The balls in the metric d take the following form:

$$r \ge 4^q+1 \implies B(0,r)=X, \tag{102}$$

$$\exists j \in \{1,\ldots,q-1\},\ 4^j+1 \le r < 4^{j+1} \implies B(0,r)=X_q \times V_j, \tag{103}$$

$$\exists j \in \{1,\ldots,q-1\},\ 4^j \le r < 4^j+2^{-m+1} \implies B(0,r)=(E_j\times V_j)\cup(X_q\times V_{j-1}), \tag{104}$$

$$\begin{aligned}\exists(j,\ell) \in \{1,\ldots,q-1\}\times\{1,\ldots,m-1\},\ & 4^j+2^{-\ell} \le r < 4^j+2^{-\ell+1}\\ & \implies B(0,r)=((E_j+W_{-\ell})\times V_j)\cup(X_q\times V_{j-1}),\end{aligned} \tag{105}$$

$$1 \le r < 4 \implies B(0,r)=X_q\times\{0\}, \tag{106}$$

$$\exists \ell \in \{1,\ldots,m\},\ 2^{-\ell} \le r < 2^{-\ell+1} \implies B(0,r)=W_{-j}\times\{0\}. \tag{107}$$

We shall first of all prove that (X,d,μ) is doubling with constant $2q \le K$. For $r \ge 4$ take $j \in \{1,2,\ldots\}$ such that $4^j \le r < 4^{j+1}$. If, in addition, $4^j+1 \le r < 4^{j+1}$ then since $2r < 4^{j+2}$, it follows from (103), (104), (105) that $B(x,2r) \subseteq X_q \times V_{j+1}$, implying that

$$\mu(B(0,2r)) \le \mu(X_q\times V_{j+1}) = q^{j+1}\mu(X_q) \le q\cdot\mu(X_q\times V_j) \stackrel{(103)}{=} q\cdot\mu(B(0,r)). \tag{108}$$

On the other hand, if $4^j \le r < 4^j+1$ then $4^j+1 \le 2r < 4^{j+1}$. Note that (104), (105) imply that

$$E_j \times V_j \subseteq B(0,r), \tag{109}$$

and therefore

$$\mu(B(0,2r)) \stackrel{(103)}{=} \mu(X_q\times V_j) = q^j\mu(X_q) \stackrel{(85)}{\le} 2q^{j+1}\mu(E_j) = 2q\mu(E_j\times V_j) \stackrel{(109)}{\le} 2q\mu(B(x,r)).$$

Similarly, using (106), (107), also for $0<r<4$ we have $\mu(B(0,2r)) \le q\mu(B(0,r))$. Thus (X,d,μ) is doubling with constant $2q$, as claimed.

Now, from (87) we can find $f_q : X_q \to \mathbb{R}_+$ with norm $\|f_q\|_{L_1(X_q)} = 1$ and $\lambda > 0$ such that

$$\mu(M_q f_q > \lambda) > \frac{q}{2\lambda}. \tag{110}$$

We extend this function f_q to a function $f : X \to \mathbb{R}_+$ defined by $f(x,y) \stackrel{\text{def}}{=} f_q(x)$ for $x \in X_q$ and $y \in \mathbb{F}_q^q$. Thus

$$\|f\|_{L_1(X)} = |\mathbb{F}_q^q| \cdot \|f_q\|_{L_1(X_q)} = q^q. \tag{111}$$

We shall next compute $M_{2^{\mathbb{Z}}} f(x,y)$ for $(x,y) \in X_q \times \mathbb{F}_q^q = X$. Actually, for very minor technical reasons we need to consider the slight variant

$$M^{\varepsilon}_{2^{\mathbb{Z}}} f(x,y) = \sup_{r \in 2^{\mathbb{Z}}} \frac{1}{\mu(B((x,y),(1+\varepsilon)r))} \int_{B((x,y),(1+\varepsilon)r)} |f(x,y)|\, d\mu(x,y) \tag{112}$$

for some small $\varepsilon > 0$, but this clearly will not make a difference since we can rescale the metric by $1+\varepsilon$.

Observe that for any $1 \le j \le q$ we have

$$M^{\varepsilon}_{2^{\mathbb{Z}}} f(x,y) \ge \frac{1}{\mu\left(B(0,(1+\varepsilon)4^j)\right)} \sum_{(x',y') \in B(0,(1+\varepsilon)4^j)} f_q(x+x').$$

Note that if $0 < \varepsilon < 4^{-q} \cdot 2^{-m+1}$ then it follows from (104) that

$$\mu\left(B(0,(1+\varepsilon)4^j)\right) \le \mu(E_j \times V_j) + \mu(X_q \times V_{j-1}) \stackrel{(85)}{\le} \frac{2}{q}\mu(X_q)q^j + \mu(X_q)q^{j-1} = 3q^{j-1}\mu(X_q).$$

Using the inclusion $B(0,(1+\varepsilon)4^j) \supseteq E_j \times V_j$, which trivially follows from (104), we conclude that

$$M^{\varepsilon}_{2^{\mathbb{Z}}} f(x,y) \ge \frac{1}{3q^{j-1}\mu(X_q)} \sum_{x' \in E_j} \sum_{y' \in V_j} f_q(x+x').$$

Hence, in combination with (98) and (85), we get the bound

$$M^{\varepsilon}_{2^{\mathbb{Z}}} f(x,y) \ge \frac{1}{6\mu(E_j)} \sum_{x' \in E_j} f_q(x+x').$$

Taking the supremum over all j we conclude the pointwise estimate

$$M^{\varepsilon}_{2^{\mathbb{Z}}} f(x,y) \ge \frac{1}{6} M_q f_q(x).$$

In particular we have

$$\mu\left(M^{\varepsilon}_{2^{\mathbb{Z}}} f > \frac{1}{6}\lambda\right) \ge |\mathbb{F}_q^q| \mu\left(M_q f_q > \lambda\right) > q^q \cdot \frac{q}{2\lambda}.$$

Recalling (111) we thus see that

$$\|M^{\varepsilon}_{2^{\mathbb{Z}}}\|_{L_1(X) \to L_{1,\infty}(X)} \ge \frac{1}{12} q \ge \frac{K}{48},$$

yielding (19).

The only remaining task is to establish the L_p bounds $\|M\|_{L_p(X) \to L_p(X)} \lesssim_p 1$, for $p > 1$. To do this let's examine what equations (102)–(107) say about the measures of the balls $B(0,r)$ appearing in the definition of the maximal function M. For $r < 4$, the balls all take the form $W_{-j} \times \{0\}$ for some $-m \le -j \le 0$. For $4^j \le r < 4^j + 2^{-m+1}$ for some $1 \le j \le q$, the ball

$B(0,r)$ is equal to the union of the two sets $E_j \times V_j$ and $X_q \times V_{j-1}$, which have the same measure up to a universal factor thanks to (85), (98). For $4^j + 2^{-m+1} \le r < 4^{j+1}$, we see that the ball $B(0,r)$ lies between $(E_j + W_{-m+1}) \times V_j$ and $X_q \times V_j$, and so thanks to (88) has measure comparable to $X_q \times V_q$. Putting all this together, we obtain the pointwise bound

$$Mg(x,y) \lesssim \max_{-m \le -j \le 0} \frac{1}{\mu(W_{-j})} \sum_{x' \in x + W_{-j}} |g(x',y)| + \max_{0 \le j \le q} \frac{1}{\mu(X_q)\mu(V_j)} \sum_{x' \in X_q} \sum_{y' \in y + V_j} |g(x',y')| \\ + \max_{0 \le j \le q} \frac{1}{\mu(E_j)\mu(V_j)} \sum_{x' \in x + E_j} \sum_{y' \in y + V_j} |g(x',y')| \quad (113)$$

for all functions $g : X \to \mathbb{R}$.

If we let $\mathscr{B}_{-j}$, for $-m \le -j \le 0$, be the σ-algebra on X generated by the cosets of $W_{-j} \times \{0\}$, we have

$$\max_{-m \le -j \le 0} \frac{1}{\mu(W_{-j})} \sum_{x' \in x + W_{-j}} |g(x',y)| = \max_{-m \le -j \le 0} \mathbb{E}\left[|g| \big| \mathscr{B}_{-j}\right](x,y), \quad (114)$$

where $\mathbb{E}\left[|g| \big| \mathscr{B}_{-j}\right]$ denotes the conditional expectation of $|g|$ with respect to the σ-algebra $\mathscr{B}_{-j}$. Applying Doob's maximal inequality (Proposition 2.1), we thus see that this expression is bounded on L_p, i.e.,

$$\left(\int_X \left| \sup_{-m \le -j \le 0} \frac{1}{\mu(W_{-j})} \sum_{x' \in x + W_{-j}} |g(x',y)| \right|^p d\mu(x,y) \right)^{1/p} \le \frac{p}{p-1} \|g\|_{L_p(X)}. \quad (115)$$

A similar argument disposes of the second term in (113), i.e.,

$$\left(\int_X \left| \max_{0 \le j \le q} \frac{1}{\mu(X_q)\mu(V_j)} \sum_{x' \in X_q} \sum_{y' \in y + V_j} |g(x',y')| \right|^p d\mu(x,y) \right)^{1/p} \le \frac{p}{p-1} \|g\|_{L_p(X)}. \quad (116)$$

By combining (115) and (116) with (113), we see that it suffices to establish the bound

$$\left(\int_X \left| \max_{0 \le j \le q} \frac{1}{\mu(E_j)\mu(V_j)} \sum_{x' \in x + E_j} \sum_{y' \in y + V_j} |g(x',y')| \right|^p d\mu(x,y) \right)^{1/p} \lesssim \left(\frac{p}{p-1} \right)^3 \|g\|_{L_p(X)}. \quad (117)$$

We can bound the left-hand side of (117) by

$$\left(\int_X \left| \max_{0 \le j \le q} \frac{1}{\mu(V_j)} \sum_{y' \in y + V_j} G(x,y') \right|^p d\mu(x,y) \right)^{1/p},$$

where

$$G(x,y') \stackrel{\mathrm{def}}{=} \max_{0 \le j \le q} \frac{1}{\mu(E_j)} \sum_{x' \in x + E_j} |g(x',y')|.$$

Applying Doob's maximal inequality again, we thus reduce to showing that

$$\|G\|_{L_p(X)} \lesssim \left(\frac{p}{p-1} \right)^2 \|g\|_{L_p(X)}.$$

But this follows from (86) (and Fubini's theorem). The proof of Theorem 1.3 is complete. $\square$

6.4. **The Ahlfors-David regular example.** Now we prove Theorem 1.4. Once again we may take n to be large, as the claim is easy for bounded n (e.g., one could take the usual Hardy-Littlewood maximal function on $\mathbb{R}^n$).

The heart of our construction is the following lemma:

Lemma 6.1. *There exists a finite Abelian group X, equipped with counting measure μ and an invariant metric d_X, with the following properties:*

(1) The are integers $a<b$ such that for all $x,y\in X$ we have $d_X(x,y)\in\{0\}\cup\{3^{j/n}\}_{j=a}^{b}$.
(2) For all $r\in[3^{a/n},3^{b/n}]$ and all $x\in X$ we have
$$3^{-a}r^n\le \mu(B(x,r))\le 3^{-a+3}r^n. \tag{118}$$
(3) $\|M\|_{L_1(X)\to L_{1,\infty}(X)}\gtrsim n\log n$.
(4) $\|M_{2^{\mathbb{Z}}}\|_{L_1(X)\to L_{1,\infty}(X)}\gtrsim \log n$.
(5) For all $1<p\le\infty$ we have $\|M\|_{L_p(X)\to L_p(X)}\lesssim_p 1$.

Proof of Theorem 1.4 assuming Lemma 6.1. In what follows $\mathbb{F}_3$ denotes the field of size 3. Let Y be the subspace of $\mathbb{F}_3^{\aleph_0}$ consisting of all finitely supported vectors, equipped with the counting measure ν. For $(y_1,y_2,\ldots)\in Y$ let $j(y)$ denote the largest $j\in\mathbb{N}$ such that $y_j\neq 0$. If $y=0$ we set $j(y)=-\infty$. For $y,y'\in Y$ define
$$\rho_Y(y,y')\stackrel{\mathrm{def}}{=}3^{b/n}\cdot 3^{j(y-y')/n}.$$
Then ρ_Y is an invariant ultrametric on Y, satisfying $\rho_Y(y,y')\in\{0\}\cup\{3^{(b+j)/n}\}_{j=1}^{\infty}$ for all $y,y'\in Y$. Let $Y_j\subseteq Y$ denote the set of vectors whose support is contained in the first j coordinates. Thus Y_j is a subspace of Y and $Y_j=B_{\rho_Y}\left(0,3^{(b+j)/n}\right)$. Since $\nu(Y_j)=3^j$, it follows that for all $r\ge 3^{b/n}$ and $y\in Y$ we have
$$3^{-b-1}r^n\le\nu\left(B_{\rho_Y}(y,r)\right)\le 3^{-b}r^n. \tag{119}$$

Next, we let Z denote the set $\mathbb{F}_3^{\aleph_0}$, and let τ denote the countable product of the normalized counting measure on $\mathbb{F}_3$. Thus τ is an invariant probability measure on Z. For $k\in\mathbb{N}$ let Z_k be the subspace of Z consisting of $(z_1,z_2,\ldots)\in Z$ with $z_1=z_2=\ldots=z_k=0$ (we shall also use the convention $Z_0=Z$). Thus $\tau(Z_k)=3^{-k}$. For $z\in Z$ let $k(z)$ denote the largest integer $k\ge 0$ such that $z\in Z_k$ (with the convention $k(0)=\infty$). For $z,z'\in Z$ define
$$\rho_Z(z,z')\stackrel{\mathrm{def}}{=}3^{(a-1)/n}\cdot 3^{-k(z-z')/n}.$$
Then ρ_Z is an invariant ultrametric on Z, satisfying $\rho_Z(z,z')\in\{0\}\cup\{3^{(a-j)/n}\}_{j=1}^{\infty}$ for all $z,z'\in Y$. It follows from the definitions that for all $k\ge 0$ we have $B_{\rho_Z}\left(0,3^{(a-k-1)/n}\right)=Z_k$. Let σ be the invariant measure on Z given by $\sigma=3^{a-1}\tau$. Thus for all $r\le 3^{a/n}$ we have,
$$\frac{1}{3}r^n\le\sigma\left(B_{\rho_Z}(y,r)\right)\le 3r^n. \tag{120}$$

We shall now let G be the Abelian group $Z\times X\times Y$, equipped with the ℓ_∞ product metric $d_G\left((z,x,y),(z',x',y')\right)=\max\left\{\rho_Z(z,z'),d(x,x'),\rho_Y(y,y')\right\}$. We shall also equip G with the product measure $\mu_G=\sigma\times\mu\times\nu$.

The balls in G are given by $B_{d_G}(0,r)=B_{\rho_Z}(0,r)\times B_{d_X}(0,r)\times B_{\rho_Y}(0,r)$. If $r\ge 3^{b/n}$ then $B_{d_X}(0,r)=X$, and thus by (118) we have $\mu\left(B_{d_X}(0,r)\right)\in\left[3^{b-a},3^{b-a+3}\right]$. Similarly,

for $r \geq 3^{b/n}$ we have $B_{\rho_Z}(0,r) = Z$, and thus $\sigma\left(B_{\rho_Z}(0,r)\right) = 3^{a-1}$. It therefore follows from (119) that

$$r \geq 3^{b/n} \implies \mu_G\left(B_{d_G}(0,r)\right) \in \left[\frac{1}{9}r^n, 9r^n\right]. \tag{121}$$

If $3^{a/n} \leq r < 3^{b/n}$, then $B_{\rho_Y}(0,r) = \{0\}$, and hence $\nu\left(B_{\rho_Y}(0,r)\right) = 1$. As before, we also have in this case $\sigma\left(B_{\rho_Z}(0,r)\right) = 3^{a-1}$, and by (118), $\mu\left(B_{d_X}(0,r)\right) \in [3^{-a}r^n, 3^{-a+3}r^n]$. Thus,

$$3^{a/n} \leq r < 3^{b/n} \implies \mu_G\left(B_{d_G}(0,r)\right) \in \left[\frac{1}{3}r^n, 9r^n\right]. \tag{122}$$

Finally, for $r < 3^{a/n}$ we have $B_{\rho_Y}(0,r) = \{0\}$ and $B_{d_X}(0,r) = \{0\}$ and so $\nu\left(B_{\rho_Y}(0,r)\right) = \mu\left(B_{d_X}(0,r)\right) = 1$. In combination with (120), we see that

$$r < 3^{a/n} \implies \mu_G\left(B_{d_G}(0,r)\right) \in \left[\frac{1}{3}r^n, 3r^n\right]. \tag{123}$$

Inequalities (121), (122), (123) show that the metric measure space (G, d_G, μ_G) is Ahlfors-David n-regular.

It remains to prove the estimates (20), (21), (22). By assertion (3) of Lemma 6.1 we can find $f : X \to \mathbb{R}_+$ with $\|f\|_{L_1(X)} = 1$, and $\lambda > 0$, such that

$$\mu\left(Mf > \lambda\right) \gtrsim \frac{n \log n}{\lambda}. \tag{124}$$

Define a function $g : G \to \mathbb{R}_+$ by $g(z,x,y) = f(x)\mathbf{1}_{\{y=0\}}$. Then $\|g\|_{L_1(G)} = \sigma(Z) = 3^{a-1}$. Moreover, we have the pointwise estimate

$$\begin{aligned} Mg(x,y,z) \geq \sup_{3^{a/n} \leq r \leq 3^{b/n}} & \frac{\int_{B_{\rho_Z}(0,r)\times B_{d_X}(0,r)\times B_{\rho_Y}(0,r)} f(x+x')\mathbf{1}_{\{z+z'=0\}} d\mu_G(z',x',y')}{\mu_G\left(B_{\rho_Z}(0,r)\times B_{d_X}(0,r)\times B_{\rho_Y}(0,r)\right)} \\ & = \left(Mf(x)\right)\mathbf{1}_{\{z=0\}}, \end{aligned}$$

where we used the fact that $B_{\rho_Y}(0,r) = \{0\}$ for $r \leq 3^{b/n}$. Thus by Fubini's theorem,

$$\mu_G\left(Mg > \lambda\right) \geq \sigma(Z)\mu\left(Mf > \lambda\right) \overset{(124)}{\gtrsim} 3^{a-1}\frac{n \log n}{\lambda} = \frac{n \log n}{\lambda}\|g\|_{L_1(G)}.$$

This proves (20); the proof of (21) is identical. To prove (22) take an non-negative $h \in L_p(G)$, and observe the pointwise bound

$$\begin{aligned} Mh(z,x,y) \ \leq\ & \sup_{r<3^{a/n}} \frac{\int_{B_{\rho_Z}(0,r)} h(z+z',x,y)d\sigma(z')}{\sigma\left(B_{\rho_Z}(0,r)\right)} && (125) \\ & + \sup_{3^{a/n}\leq r\leq 3^{b/n}} \frac{\int_{Z\times B_{d_X}(0,r)} h(z',x+x',y)d\sigma(z')d\mu(x')}{2^{a-1}\mu\left(B_{d_X}(0,r)\right)} && (126) \\ & + \sup_{r>3^{b/n}} \frac{\int_{Z\times X\times B_{\rho_Y}(0,r)} h(z',x',y+y')d\sigma(z')d\mu(x')d\nu(y')}{2^{a-1}\cdot 3^{b-a}\nu\left(B_{\rho_Y}(0,r)\right)}, && (127) \end{aligned}$$

where in the denominator of (127) we used the fact that $\mu(x) \geq 3^{b-a}$.

Since ρ_Z is an ultrametric, Doob's maximal inequality implies that for all $x \in X$ and $y \in Y$ we have,

$$\int_Z \left(\sup_{r<3^{a/n}} \frac{\int_{B_{\rho_Z}(0,r)} h(z+z',x,y)d\sigma(z')}{\sigma\left(B_{\rho_Z}(0,r)\right)} \right)^p d\sigma(z) \lesssim_p \int_Z h(z,x,y)^p d\sigma(z).$$

Thus, by Fubini's theorem, the $L_p(G)$ norm of the term in (125) is $\lesssim_p \|h\|_{L_p(G)}$. A similar argument shows that the $L_p(G)$ norm of the term in (127) is $\lesssim_p \|h\|_{L_p(G)}$. Finally, using assertion (5) of Lemma 6.1, we get the same bound for the term in (126), proving (22). □

Proof of Lemma 6.1. Let $q = 3^k$ be a power of three between $\frac{1}{3}n \log n$ and $\frac{1}{9}n \log n$. We invoke Proposition 6.2 to create a vector space $X_q = \mathbb{F}_q^m$ over a finite field $\mathbb{F}_q$ with counting measure μ, together with sets $E_1, \ldots, E_q$ obeying the properties stated in Proposition 6.2; in particular

$$m \lesssim \sqrt{n \log n}. \tag{128}$$

Note that $\mathbb{F}_q$ can itself be viewed as a vector space over the field $\mathbb{F}_3$ of three elements, and thus X_q is a vector space over $\mathbb{F}_3$ of dimension

$$M \stackrel{\mathrm{def}}{=} mk = m \log_3 q \lesssim n^{1/2} (\log n)^{3/2}. \tag{129}$$

As in Section 6.3, the idea is to take a Cartesian product of X_q with another vector space, and try to create balls which resemble the product of a set E_j with a subspace. Some care is however required in order to make the construction compatible with both the constraint (8) and the triangle inequality.

Analogously to the arguments in Section 6.3, we shall need a flag

$$\{0\} = W_{-M} \subseteq W_{-M+1} \subseteq \ldots \subseteq W_0 = X_q$$

of vector spaces over $\mathbb{F}_3$ in X_q, so that $\mu(W_{-j}) = 3^{-j}\mu(X_q)$ for all $-M \le -j \le 0$. (We will not use (88) or the space W_{-m+1} in Proposition 6.2, so there is no collision of notation here.)

Our space shall be $X \stackrel{\mathrm{def}}{=} X_q \times \mathbb{F}_3^q$, with counting measure μ. We shall need a flag

$$\{0\} = V_0 \subseteq V_1 \subseteq \ldots \subseteq V_q = \mathbb{F}_3^q$$

in $\mathbb{F}_3^q$, with $\mu(V_j) = 3^j$.

For every integer $-M \le j \le q$, we define the set $B_j \subseteq X = X_q \times \mathbb{F}_3^q$ as follows:

- If $-M \le j \le 0$, we set $B_j \stackrel{\mathrm{def}}{=} W_j \times \{0\}$.
- If $1 \le j \le q$, we set

$$B_j \stackrel{\mathrm{def}}{=} (X_q \times V_j) \bigcup \left(\bigcup_{\ell=1}^{\min\{j+k,q\}} (E_\ell \times V_\ell) \right). \tag{130}$$

The B_j are symmetric and nested, with

$$\{0\} = B_{-M} \subseteq B_{-M+1} \subseteq \ldots \subseteq B_q = X. \tag{131}$$

We define a function $d : X \times X \to \mathbb{R}_+$ by setting $d(x,x) = 0$ for all $x \in X$, and

$$d(x,y) \stackrel{\mathrm{def}}{=} \min\left\{3^{j/n} : x - y \in B_j\right\}, \tag{132}$$

for all distinct $x, y \in X$. Thus d takes values in $\{0\} \cup \left\{3^{j/n} :\ -M+1 \le j \le q\right\}$. The first assertion of Lemma 6.1 therefore holds with $a = -M+1$ and $b = q$.

Claim 6.2. *d is a translation-invariant metric on X.*

Proof. The translation-invariance, non-degeneracy, and symmetry properties of d are obvious (symmetry follows from the symmetry of E_j). The only non-trivial task is to verify the triangle inequality. By construction, it will suffice to show that $x + x' \in B_{j''-1}$ whenever $x \in B_j$, $x' \in B_{j'}$, and $-M < j, j', j'' \le q$ are such that

$$3^{j''/n} > 3^{j/n} + 3^{j'/n}. \tag{133}$$

By symmetry we may assume that $j \le j'$. It follows from (133) that provided n is large enough,

$$3^{j''/n} > 3^{j'/n}\left(1 + 3^{-(M+q-1)/n}\right) \ge 3^{j'/n}\left(1 + 3^{-\frac{1}{2}\log n}\right) \ge 3^{(j'+k)/n}, \tag{134}$$

where we used the fact that $q \le \frac{1}{3} n \log n$, while $M \lesssim n^{1/2}(\log n)^{3/2}$ and $k = \log_3 q \lesssim \log n$.

It follows from (134) that

$$j'' > j' + k. \tag{135}$$

If $j' \le 0$, then we have $B_j + B_{j'} = B_{j'}$, so $x + x' \in B_{j'} \subseteq B_{j''-k} \subseteq B_{j''-1}$, as required. Assume therefore that $j' \ge 1$. Then $B_j \subseteq B_{j'} \subseteq X_q \times V_{\min\{j'+k,q\}}$, and hence $x + x' \in X_q \times V_{\min\{j'+k,q\}}$. On the other hand, we will have $X_q \times V_{\min\{j'+k,q\}} \subseteq B_{j''-1}$ as soon as $\min\{j'+k, q\} < j''$. Since $j'' \le q$, it follows from (135) that $j' + k < q$. Hence, using (135) once more, we see that $\min\{j' + k, q\} = j' + k < j''$, as required. □

Claim 6.3. *For all $r \in \left[3^{-(M-1)/n}, 3^{q/n}\right]$ and all $x \in X$, we have*

$$\frac{1}{3} r^n \le \frac{\mu(B(x,r))}{\mu(X_q)} \le 4r^n.$$

Proof. By translation invariance we may assume that $x = 0$. Let j be the integer such that $3^{j/n} \le r < 3^{(j+1)/n}$. Then $B(0,r) = B_j$. If $j \le 0$ then $B_j = W_j \times \{0\}$, and hence

$$\frac{\mu(B(x,r))}{\mu(X_q)} = \frac{\mu(B_j)}{\mu(X_q)} = 3^j \in \left[\frac{1}{3} r^n, r^n\right]. \tag{136}$$

If $j \ge 1$ the it follows from (130) that $B_j \supseteq X_q \times V_j$, and hence

$$\frac{\mu(B(0,r))}{\mu(X_q)} = \frac{\mu(B_j)}{\mu(X_q)} \ge \mu(V_j) = 3^j \ge \frac{1}{3} r^n. \tag{137}$$

At the same time, it follows from (130) that

$$\begin{aligned}\frac{\mu(B(0,r))}{\mu(X_q)} = \frac{\mu(B_j)}{\mu(X_q)} \le 3^j + \sum_{\ell=1}^{\min\{j+k,q\}} \frac{\mu(E_\ell)}{\mu(X_q)} \mu(V_\ell) &\overset{(85)}{\le} 3^j + \frac{2}{q} \sum_{\ell=1}^{\min\{j+k,q\}} 3^j \\ &\le 3^j + \frac{3}{q} \cdot 3^{\min\{j+k,q\}} = 3^j + \frac{3}{3^k} \cdot 3^{\min\{j+k,q\}} = 4 \cdot 3^j \le 4r^n, \end{aligned} \tag{138}$$

as required. □

Claim 6.3 implies the second assertion of Lemma 6.1, since $\mu(X_q) = 3^M = 3^{-a+1}$.

We shall now prove the third assertion of Lemma 6.1. Since the balls $B(x,r)$ in X take the form $x + B_j$ for some j, we have

$$Mf(x) = \max_{-M \le j \le q} \frac{1}{\mu(B_j)} \sum_{y \in B_j} |f(x+y)|.$$

¿From (87) we can find $f_q : X_q \to \mathbb{R}_+$ with $\|f_q\|_{L_1(X_q)} = 1$, and $\lambda > 0$, such that
$$\mu\left(M_q f_q > \lambda\right) > \frac{q}{2\lambda}. \tag{139}$$
We extend this function f_q to a function $f : X \to \mathbb{R}_+$ defined by $f(x,y) \stackrel{\mathrm{def}}{=} f_q(x)$ for $x \in X_q$ and $y \in \mathbb{F}_3^q$. Thus,
$$\|f\|_{L_1(X)} = 3^q. \tag{140}$$
Observe that for $1 \le j \le q-k$ we have,
$$\mu(B_j) \stackrel{(138)}{\le} 4\mu(X_q)3^j \stackrel{(85)}{\le} 8q\mu(E_{j+k})\mu(V_j) = 8\mu(E_{j+k})\mu(V_{j+k}).$$
Hence, for all $(x,x') \in X$ we have,
$$Mf(x,x') \ge \max_{1\le j\le q-k} \frac{1}{8\mu(V_{j+k})\mu(E_{j+k})} \sum_{(y,y')\in B_j} |f_q(x+y)|.$$
Since B_j contains $E_{j+k} \times V_{j+k}$ for $1 \le j \le q-k$, we conclude that
$$\begin{aligned} Mf(x,x') &\ge \max_{k+1\le \ell\le q} \frac{1}{8\mu(E_\ell)} \sum_{y\in E_\ell} |f_q(x+y)| \\ &\ge \frac{1}{8}\left(\max_{1\le j\le q} \frac{1}{\mu(E_j)} \sum_{y\in E_j} |f_q(x+y)| - \sum_{j=1}^{k} \frac{1}{\mu(E_j)} \sum_{y\in E_j} |f_q(x+y)|\right). \end{aligned} \tag{141}$$
Denote $g : X \to \mathbb{R}$ by $g(x,x') = \sum_{j=1}^{k} \frac{1}{\mu(E_j)} \sum_{y\in E_j} |f_q(x+y)|$. Then
$$\|g\|_{L_1(X)} \le k3^q \|f_q\|_{L_1(X_q)} \stackrel{(140)}{=} \|f\|_{L_1(X)} \log_3 q. \tag{142}$$
It follows from (141) that we have the pointwise bound $M_q f_q(x) \le 8Mf(x,x') + g(x,x')$. Thus,
$$\begin{aligned} \frac{q\|f\|_{L_1(X)}}{2\lambda} &\stackrel{(139)\wedge(140)}{\le} \mu\left((x,x') \in X :\ M_q f_q(x) > \lambda\right) \\ &\le \mu\left(8Mf + g > \lambda\right) \\ &\le \mu\left(Mf > \frac{\lambda}{16}\right) + \mu\left(g > \frac{\lambda}{2}\right) \\ &\le \mu\left(Mf > \frac{\lambda}{16}\right) + \frac{2\|g\|_{L_1(X)}}{\lambda} \\ &\stackrel{(142)}{\le} \mu\left(Mf > \frac{\lambda}{16}\right) + \frac{2\log_3 q\|f\|_{L_1(X)}}{\lambda}. \end{aligned}$$
Hence,
$$\|M\|_{L_1(X)\to L_{1,\infty}(X)} \gtrsim q \gtrsim n\log n, \tag{143}$$
which gives the third assertion of Lemma 6.1.

A similar argument (requiring a closer inspection of the details of Proposition 6.2) can be used to give the fourth assertion of Lemma 6.1; alternatively, one can use (143) and the pigeonhole principle to show that a dilated version $M_{r\cdot 2^{\mathbb{Z}}}$ of the lacunary maximal function

has weak $(1,1)$ norm $\gtrsim \log n$ for some $r>0$, and then rescale the metric. We omit the details.

It remains to verify the L_p bound in assertion (5) of Lemma 6.1, i.e., to show for all $f\in L_p(X)$ we have

$$\left\|\max_{-M\le j\le q}\frac{1}{\mu(B_j)}\sum_{y\in B_j}|f(x+y)|\right\|_{L_p(X)}\lesssim_p \|f\|_{L_p(X)}.$$

The contribution of the case $-M\le j\le 0$ can be handled by Doob's maximal inequality as in Section 6.3, so we need only consider the case $1\le j\le q$. Using (**??**) and the definition of B_j, we soon verify the pointwise estimate

$$\max_{1\le j\le q}\frac{1}{\mu(B_j)}\sum_{y\in B_j}|f(x+y)|\lesssim \max_{1\le i\le q}\frac{1}{\mu(X_q\times V_i)}\sum_{y\in X_q\times V_i}|f(x+y)| + \max_{1\le i\le q}\frac{1}{\mu(E_i\times V_i)}\sum_{y\in E_i\times V_i}|f(x+y)|. \tag{144}$$

Indeed, denote

$$h(x)=\max_{1\le i\le q}\frac{1}{\mu(X_q\times V_i)}\sum_{y\in X_q\times V_i}|f(x+y)|+\max_{1\le i\le q}\frac{1}{\mu(E_i\times V_i)}\sum_{y\in E_i\times V_i}|f(x+y)|.$$

Then for all $1\le j\le q$,

$$\frac{1}{\mu(B_j)}\sum_{y\in B_j}|f(x+y)| \overset{(130)}{\lesssim} \frac{\sum_{y\in X_q\times V_j}|f(x+y)|+\sum_{\ell=1}^{\min\{j+k,q\}}\sum_{y\in E_\ell\times V_\ell}|f(x+y)|}{\mu(B_j)}$$
$$\overset{(137)}{\le}\frac{h(x)\mu\left(X_q\times V_j\right)+\sum_{\ell=1}^{\min\{j+k,q\}}h(x)\mu\left(E_\ell\times V_\ell\right)}{3^j\mu(X_q)}\overset{(138)}{\le} 4h(x),$$

proving (144).

The fact that the first term in the right-hand side of (144) is bounded in $L_p(X)$ again follows from Doob's maximal inequality, while the $L_p(X)$ boundedness of the second term in the right-hand side of (144) follows from (86), Doob's maximal inequality and a Fubini argument, as in Section 6.3. The proof of Theorem 1.4 is now complete. □

References

[1] J. M. Aldaz. The weak type $(1,1)$ bounds for the maximal function associated to cubes grow to infinity with the dimension. Preprint. Available at http://arxiv.org/abs/0805.1565, 2008.

[2] N. Alon, R. M. Karp, D. Peleg, and D. West. A graph-theoretic game and its application to the k-server problem. *SIAM J. Comput.*, 24(1):78–100, 1995.

[3] G. Aubrun. Maximal inequality for high-dimensional cubes. *Confluentes Mathematici*, 1:169–179, 2009.

[4] Y. Bartal. Probabilistic approximation of metric spaces and its algorithmic applications. In *37th Annual Symposium on Foundations of Computer Science (Burlington, VT, 1996)*, pages 184–193. IEEE Comput. Soc. Press, Los Alamitos, CA, 1996.

[5] Y. Bartal. On approximating arbitrary metrices by tree metrics. In *STOC '98 (Dallas, TX)*, pages 161–168. ACM, New York, 1999.

[6] A. S. Besicovitch. A general form of the covering principle and relative differentiation of additive functions. *Proc. Cambridge Philos. Soc.*, 41:103–110, 1945.

[7] A. S. Besicovitch. A general form of the covering principle and relative differentiation of additive functions. II. *Proc. Cambridge Philos. Soc.*, 42:1–10, 1946.

[8] J. Bourgain. On high-dimensional maximal functions associated to convex bodies. *Amer. J. Math.*, 108(6):1467–1476, 1986.

[9] J. Bourgain. On the L^p-bounds for maximal functions associated to convex bodies in $\mathbf{R}^n$. *Israel J. Math.*, 54(3):257–265, 1986.

[10] J. Bourgain. On dimension free maximal inequalities for convex symmetric bodies in $\mathbf{R}^n$. In *Geometrical aspects of functional analysis (1985/86)*, volume 1267 of *Lecture Notes in Math.*, pages 168–176. Springer, Berlin, 1987.

[11] J. Bourgain. Remarks on the extension of Lipschitz maps defined on discrete sets and uniform homeomorphisms. In *Geometrical aspects of functional analysis (1985/86)*, volume 1267 of *Lecture Notes in Math.*, pages 157–167. Springer, Berlin, 1987.

[12] G. Calinescu, H. Karloff, and Y. Rabani. Approximation algorithms for the 0-extension problem. *SIAM J. Comput.*, 34(2):358–372 (electronic), 2004/05.

[13] A. Carbery. An almost-orthogonality principle with applications to maximal functions associated to convex bodies. *Bull. Amer. Math. Soc. (N.S.)*, 14(2):269–273, 1986.

[14] J. Cheeger. Differentiability of Lipschitz functions on metric measure spaces. *Geom. Funct. Anal.*, 9(3):428–517, 1999.

[15] J. Cheeger, B. Kleiner, and A. Naor. Compression bounds for Lipschitz maps from the Heisenberg group to L_1. Preprint. Available at http://arxiv.org/abs/0910.2026, 2009.

[16] M. Christ. A $T(b)$ theorem with remarks on analytic capacity and the Cauchy integral. *Colloq. Math.*, 60/61(2):601–628, 1990.

[17] R. R. Coifman and G. Weiss. *Analyse harmonique non-commutative sur certains espaces homogènes.* Lecture Notes in Mathematics, Vol. 242. Springer-Verlag, Berlin, 1971. Étude de certaines intégrales singulières.

[18] T. H. Colding and W. P. Minicozzi, II. Liouville theorems for harmonic sections and applications. *Comm. Pure Appl. Math.*, 51(2):113–138, 1998.

[19] J. L. Doob. *Stochastic processes*. Wiley Classics Library. John Wiley & Sons Inc., New York, 1990. Reprint of the 1953 original, A Wiley-Interscience Publication.

[20] R. Dror, S. Ganguli, and R. S. Strichartz. A search for best constants in the Hardy-Littlewood maximal theorem. *J. Fourier Anal. Appl.*, 2(5):473–486, 1996.

[21] R. Durrett. *Probability: theory and examples*. Duxbury Press, Belmont, CA, second edition, 1996.

[22] J. Fakcharoenphol, S. Rao, and K. Talwar. A tight bound on approximating arbitrary metrics by tree metrics. *J. Comput. System Sci.*, 69(3):485–497, 2004.

[23] V. Fischer. The spherical maximal function on the free two-step nilpotent Lie group. *Math. Scand.*, 99(1):99–118, 2006.

[24] Z. Füredi and P. A. Loeb. On the best constant for the Besicovitch covering theorem. *Proc. Amer. Math. Soc.*, 121(4):1063–1073, 1994.

[25] L. Grafakos, S. Montgomery-Smith, and O. Motrunich. A sharp estimate for the Hardy-Littlewood maximal function. *Studia Math.*, 134(1):57–67, 1999.

[26] O. Guédon and A. Zvavitch. Supremum of a process in terms of trees. In *Geometric aspects of functional analysis*, volume 1807 of *Lecture Notes in Math.*, pages 136–147. Springer, Berlin, 2003.

[27] J. Heinonen. *Lectures on analysis on metric spaces*. Universitext. Springer-Verlag, New York, 2001.

[28] J. R. Lee and A. Naor. Extending Lipschitz functions via random metric partitions. *Invent. Math.*, 160(1):59–95, 2005.

[29] A. J. Lemin. Isometric embedding of ultrametric (non-Archimedean) spaces in Hilbert space and Lebesgue space. In *p-adic functional analysis (Ioannina, 2000)*, volume 222 of *Lecture Notes in Pure and Appl. Math.*, pages 203–218. Dekker, New York, 2001.

[30] H.-Q. Li. Fonctions maximales centrées de Hardy-Littlewood sur les groupes de Heisenberg. *Studia Math.*, 191(1):89–100, 2009.

[31] E. Lindenstrauss. Pointwise theorems for amenable groups. *Invent. Math.*, 146(2):259–295, 2001.

[32] N. Linial and M. Saks. Low diameter graph decompositions. *Combinatorica*, 13(4):441–454, 1993.

[33] J. Luukkainen and E. Saksman. Every complete doubling metric space carries a doubling measure. *Proc. Amer. Math. Soc.*, 126(2):531–534, 1998.

[34] A. D. Melas. The best constant for the centered Hardy-Littlewood maximal inequality. *Ann. of Math. (2)*, 157(2):647–688, 2003.

[35] M. T. Menárguez and F. Soria. On the maximal operator associated to a convex body in $\mathbf{R}^n$. *Collect. Math.*, 43(3):243–251 (1993), 1992.

[36] M. Mendel and A. Naor. Maximum gradient embeddings and monotone clustering. To appear in Combinatorica. Preprint available at http://arxiv.org/abs/cs.DS/0606109, 2006.

[37] M. Mendel and A. Naor. Ramsey partitions and proximity data structures. *J. Eur. Math. Soc.*, 9(2):253–275, 2007.

[38] D. Müller and A. Seeger. Singular spherical maximal operators on a class of two step nilpotent Lie groups. *Israel J. Math.*, 141:315–340, 2004.

[39] F. Nazarov, S. Treil, and A. Volberg. The Tb-theorem on non-homogeneous spaces. *Acta Math.*, 190(2):151–239, 2003.

[40] A. Nevo and E. M. Stein. A generalization of Birkhoff's pointwise ergodic theorem. *Acta Math.*, 173(1):135–154, 1994.

[41] A. Nevo and S. Thangavelu. Pointwise ergodic theorems for radial averages on the Heisenberg group. *Adv. Math.*, 127(2):307–334, 1997.

[42] R. Rochberg and M. Taibleson. Factorization of the Green's operator and weak-type estimates for a random walk on a tree. *Publ. Mat.*, 35(1):187–207, 1991. Conference on Mathematical Analysis (El Escorial, 1989).

[43] Y. Sawano. Sharp estimates of the modified Hardy-Littlewood maximal operator on the nonhomogeneous space via covering lemmas. *Hokkaido Math. J.*, 34(2):435–458, 2005.

[44] R. Shvidkoy. A dimension-dependent maximal inequality. *Proc. Edinb. Math. Soc. (2)*, 45(3):547–548, 2002.

[45] E. M. Stein. The development of square functions in the work of A. Zygmund. *Bull. Amer. Math. Soc. (N.S.)*, 7(2):359–376, 1982.

[46] E. M. Stein. *Harmonic analysis: real-variable methods, orthogonality, and oscillatory integrals*, volume 43 of *Princeton Mathematical Series*. Princeton University Press, Princeton, NJ, 1993. With the assistance of Timothy S. Murphy, Monographs in Harmonic Analysis, III.

[47] E. M. Stein and J.-O. Strömberg. Behavior of maximal functions in $\mathbf{R}^n$ for large n. *Ark. Mat.*, 21(2):259–269, 1983.

[48] J.-O. Strömberg. Weak type L^1 estimates for maximal functions on noncompact symmetric spaces. *Ann. of Math. (2)*, 114(1):115–126, 1981.

[49] M. Talagrand. Regularity of Gaussian processes. *Acta Math.*, 159(1-2):99–149, 1987.

[50] M. Talagrand. *The generic chaining*. Springer Monographs in Mathematics. Springer-Verlag, Berlin, 2005. Upper and lower bounds of stochastic processes.

[51] T. Tao and V. Vu. *Additive combinatorics*, volume 105 of *Cambridge Studies in Advanced Mathematics*. Cambridge University Press, Cambridge, 2006.

[52] Y. Terasawa. Outer measures and weak type $(1,1)$ estimates of Hardy-Littlewood maximal operators. *J. Inequal. Appl.*, pages Art. ID 15063, 13, 2006.

[53] J. Tišer. Differentiation theorem for Gaussian measures on Hilbert space. *Trans. Amer. Math. Soc.*, 308(2):655–666, 1988.

[54] A. L. Vol′berg and S. V. Konyagin. On measures with the doubling condition. *Izv. Akad. Nauk SSSR Ser. Mat.*, 51(3):666–675, 1987.

[55] J.-M. Wu. Hausdorff dimension and doubling measures on metric spaces. *Proc. Amer. Math. Soc.*, 126(5):1453–1459, 1998.

[56] J. Zienkiewicz. Estimates for the Hardy-Littlewood maximal function on the Heisenberg group. *Colloq. Math.*, 103(2):199–205, 2005.

[57] A. Zygmund. *Trigonometric series. Volumes I and II combined. With a foreword by Robert Fefferman. 3rd ed.* Cambridge Mathematical Library. Cambridge: Cambridge University Press. xiii, 2002.